

# **Cosmic sound: Measuring the Universe with baryonic acoustic oscillations**

Dissertation

der Fakultät für Physik der  
Ludwig-Maximilians-Universität München

angefertigt von  
**Gert Hütsi**  
aus Tallinn (Estland)



München, den 1. März 2006

Gert Hütsi:

*Cosmic sound: Measuring the Universe with baryonic acoustic oscillations*

Dissertation der Fakultät für Physik der Ludwig-Maximilians-Universität München  
ausgeführt am Max-Planck-Institut für Astrophysik

1. Gutachter:

Prof. Dr. Rashid Sunyaev, MPA Garching

2. Gutachter:

Prof. Dr. Viatcheslav Mukhanov, LMU München

Tag der mündlichen Prüfung: 30. Mai 2006

# Contents

<b>1. Introduction</b>	<b>3</b>
1.1. Standard Model of cosmology	4
1.1.1. Homogeneous-isotropic backgrounds. Classical cosmological tests	4
1.1.2. Perturbed FLRW models	9
1.2. Introduction to acoustic oscillations	13
1.3. Cosmological parameters. Markov Chain Monte Carlo	18
1.4. Large galaxy/cluster surveys. Fast semianalytical methods for structure formation	21
1.5. In this Thesis	28
<b>2. Clustering of SZ clusters on a past light-cone: acoustic oscillations and constraints on dark energy</b>	<b>31</b>
2.1. Introduction	31
2.2. Light-cone power spectrum of galaxy clusters	33
2.2.1. Cluster power spectra from VIRGO simulations	34
2.2.2. Comparison with the analytical description: accuracy of the biasing scheme	41
2.3. SZ clusters and baryonic oscillations	43
2.3.1. SZ-selected clusters. Mass-observable relations	46
2.3.2. Accuracy of the power spectrum determination	47
2.3.3. Prospects of detecting baryonic “wiggles”. Comparison with SDSS LRG	50
2.3.4. Some remarks on SZ vs. optical cluster selection	53
2.4. Constraints on Dark Energy	54
2.4.1. 2D power spectrum on a light-cone	54
2.4.2. Parameter estimation	59
2.5. Conclusions	63
<b>3. Acoustic oscillations in the SDSS DR4 Luminous Red Galaxy sample power spectrum</b>	<b>65</b>
3.1. Introduction	65
3.2. Data	67
3.3. Power spectrum calculation	69
3.4. Power spectrum errors and covariance matrix	73
3.5. Relation to the true spectrum	75
3.6. Model spectra	78
3.7. Determination of the acoustic scale	81
3.8. Correlation function analysis	83
3.9. Comparison with the other surveys	85
3.10. Discussion and Conclusions	86

<b>4. Power spectrum of the SDSS luminous red galaxies: constraints on cosmological parameters</b>	<b>89</b>
4.1. Introduction . . . . .	89
4.2. Data . . . . .	91
4.3. Power spectrum / acoustic scale transformation . . . . .	92
4.4. Results . . . . .	99
4.4.1. $W_{\text{MAP}}$ + HST data . . . . .	99
4.4.2. Constraints from the measurement of the acoustic scale . . . . .	102
4.4.3. Constraints from the full power spectrum . . . . .	107
4.4.4. One dimensional distributions . . . . .	107
4.4.5. Most interesting constraints . . . . .	108
4.5. Discussion and Conclusions . . . . .	113
<b>5. Conclusions</b>	<b>117</b>
<b>A. Fitting formulae for the acoustic scales</b>	<b>119</b>
<b>B. Test problem</b>	<b>123</b>
<b>C. Mock catalogs</b>	<b>125</b>
<b>D. Power spectrum from the halo model</b>	<b>131</b>
<b>E. Fitting formulae for the coupling kernels</b>	<b>133</b>
<b>F. Nonlinear model fitting. Correlated data</b>	<b>135</b>
<b>G. Goodness of fit. Correlated Gaussian data</b>	<b>137</b>
<b>H. SDSS LRG power spectrum and covariance matrix</b>	<b>139</b>
<b>Bibliography</b>	<b>141</b>

# Abstract

During the last ten to fifteen years cosmology has turned from a data-starved to a data-driven science. Several key parameters of the Universe have now been measured with an accuracy better than 10%. Surprisingly, it has been found that instead of slowing down, the expansion of the Universe proceeds at an ever increasing rate. From this we infer the existence of a negative pressure component– the so-called Dark Energy (DE)– that makes up more than two thirds of the total matter-energy content of our Universe. It is generally agreed amongst cosmologists and high energy physicists that understanding the nature of the DE poses one of the biggest challenges for the modern theoretical physics.

Future cosmological datasets, being superior in both quantity and quality to currently existing data, hold the promise for unveiling many of the properties of the mysterious DE component. With ever larger datasets, as the statistical errors decrease, one needs to have a very good control over the possible systematic uncertainties. To make progress, one has to concentrate the observational effort towards the phenomena that are theoretically best understood and also least “contaminated” by complex astrophysical processes or several intervening foregrounds. Currently by far the cleanest cosmological information has been obtained through measurements of the angular temperature fluctuations of the Cosmic Microwave Background (CMB). The typical angular size of the CMB temperature fluctuations is determined by the distance the sound waves in the tightly coupled baryon-photon fluid can have traveled since the Big Bang until the epoch of recombination. A similar scale is also expected to be imprinted in the large-scale matter distribution as traced by, for instance, galaxies or galaxy clusters. Measurements of the peaks in the CMB angular power spectrum fix the physical scale of the sound horizon with a high precision. By identifying the corresponding features in the low redshift matter power spectrum one is able to put constraints on several cosmological parameters.

In this thesis we have investigated the prospects for the future wide-field SZ cluster surveys to detect the acoustic scale in the matter power spectrum, specifically concentrating on the possibilities for constraining the properties of the DE. The core part of the thesis is concerned with a power spectrum analysis of the SDSS Luminous Red Galaxy (LRG) sample. We have been able to detect acoustic features in the redshift-space power spectrum of LRGs down to scales of  $\sim 0.2 h\text{Mpc}^{-1}$ , which approximately corresponds to the seventh peak in the CMB angular spectrum. Using this power spectrum measurement along with the measured size of the sound horizon, we have carried out the maximum likelihood cosmological parameter estimation using Markov chain Monte Carlo techniques. The precise measurement of the low redshift sound horizon in combination with the CMB data has enabled us to measure, under some simplifying assumptions, the Hubble constant with a high precision:  $H_0 = 70.8^{+1.9}_{-1.8} \text{ km/s/Mpc}$ . Also we have shown that a decelerating expansion of the Universe is ruled out at more than  $5\sigma$  confidence level.



# 1. Introduction

Cosmology is the study of the origin and evolution of our Universe and as such it has fairly ambitious tasks. Over the last 10-15 years, due to the rapid development of observational cosmology, our knowledge about the Universe has increased dramatically. One can say that cosmology has turned from a data-starved to a data-driven science. Several key parameters of our Universe have been measured to an accuracy better than a few percent. Moreover, we have entered a stage where we can start making accurate tests for many of the underlying assumptions. This success has led to the establishment of the Standard Model of cosmology, often also called the “Concordance” Model, in order to avoid confusion with the Standard Model of particle physics. Although it is very successful in explaining the great body of diverse observational data (and that with the model having in its simplest form only 5-6 free parameters!), we have to be worried about the doubly occurring word “Dark”, which probably also adequately describes our current level of knowledge. According to current best estimates, approximately two-thirds of our Universe is made up of the mysterious smoothly-distributed Dark Energy (DE) component with negative pressure, about one-third is in the form of the pressureless and noninteracting Cold Dark Matter (CDM), while the familiar baryonic matter makes up less than 5% of the total density. Although we have not yet detected particles possibly making up the CDM there are a plenty of candidates provided by the various extensions of the Standard Model of particle physics. Concerning the DE the situation is much less satisfactory. According to the most popular beliefs the DE might either be caused by the famous  $\Lambda$ -term introduced by Einstein, leading to the so-called  $\Lambda$ CDM model, or by a time-varying scalar field rolling slowly down the potential and thus being dominated by the potential energy. The last models are known under the name Quintessence.

In this introductory chapter we start with a very brief review of the basics of the “concordance” cosmological model.<sup>1</sup> The rest of the introduction is devoted to a discussion of acoustic oscillations occurring in the tightly-coupled baryon-photon fluid in the early radiation-dominated Universe and the imprint they leave on the angular power spectrum of the Cosmic Microwave Background (CMB) radiation and on the spatial power spectrum of galaxies/galaxy clusters. Finally we describe some of the existing and future galaxy and galaxy cluster redshift surveys and stress the need for fast semianalytical tools for analyzing these gigantic datasets.

---

<sup>1</sup>A detailed presentation of many of the underlying ideas and results leading to the standard cosmological model can be found in several cosmology textbooks e.g. Weinberg (1972), Peebles (1980), Kolb & Turner (1990), Padmanabhan (1993), Peebles (1993), Coles & Lucchin (1995), Liddle & Lyth (2000), Dodelson (2003).

## 1.1. Standard Model of cosmology

### 1.1.1. Homogeneous-isotropic backgrounds. Classical cosmological tests

Looking at the complex patterns of the large-scale structure surrounding us, it seems at first sight pretty hopeless to find any theory capable of describing all that richness. To make any progress one certainly has to adopt several simplifying assumptions. The most important of these is the Cosmological Principle, which states that on the largest scales our Universe is homogeneous and isotropic, i.e. there are no special locations and directions, and thus it can be seen as a generalization of the Copernican principle for the whole Universe. For decades this assumption was taken for granted, due to aesthetic reasons, without any precise observational proof. Currently we have a very good test for the assumption of isotropy provided by all-sky CMB experiments like COBE and more lately by WMAP, which demonstrate that after the subtraction of the dipole caused by the peculiar motion of the Local Group, the remaining angular temperature fluctuations with respect to the mean are only of the order  $10^{-5}$ . The isotropy around all the other points, as can be expected if our location is by no means special, also implies homogeneity. If these symmetries hold, and in addition the Weyl's postulate, which states that there is a unique geodesic at each space-time point, is satisfied, the most general space-time interval can be written as:

$$ds^2 = c^2 dt^2 - R^2(t) \left[ d\chi^2 + S_k^2(\chi)(d\theta^2 + \sin^2 \theta d\phi^2) \right]. \quad (1.1)$$

This is the Robertson-Walker metric. Here

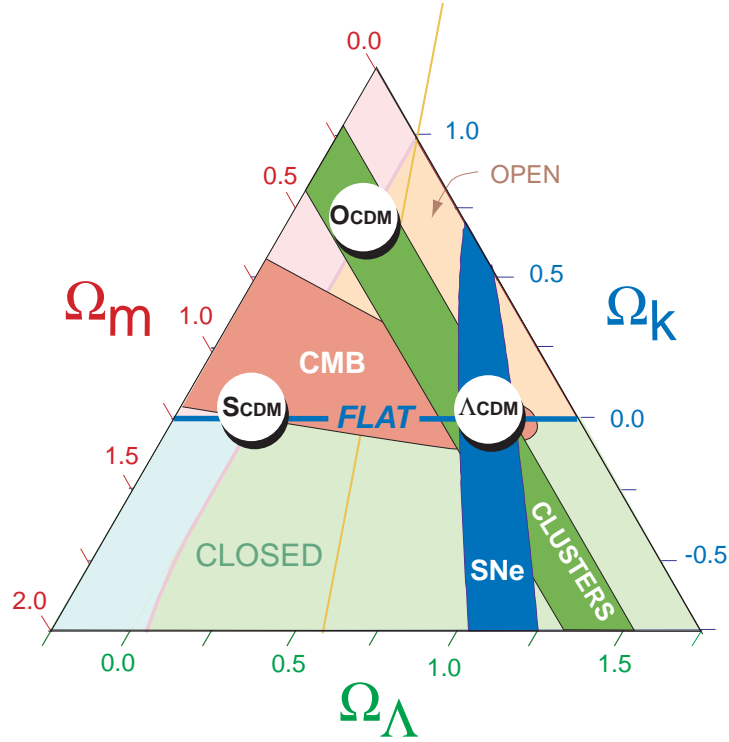
$$S_k(\chi) = \begin{cases} \sin \chi & \text{if } k = 1 \\ \chi & \text{if } k = 0 \\ \sinh \chi & \text{if } k = -1, \end{cases} \quad (1.2)$$

$t$  is the proper time measured by comoving observers, the scale factor  $R(t)$  is the radius of curvature of space-like sections,  $\chi$  is a dimensionless radial coordinate,  $\theta$  and  $\phi$  are the usual spherical coordinates, and  $k$  is known as the curvature parameter. Only the sign of the curvature is important since the general case can always be reduced to one of the above three cases by a suitable rescaling. Assuming that General Relativity gives a valid description for the space-time evolution even at cosmologically large scales, we can move on and solve the Einstein field equations to obtain the function  $R(t)$ .<sup>2</sup> To do that we also have to specify the matter-energy content of the model universe. The standard assumption here is that Universe consists of several distinct components like radiation, matter, and DE, which can be described as perfect fluids. In order to close the set of resulting equations, the so-called Friedmann equations, one also has to provide the relation between pressure and density for each of the components, i.e. one has to specify the equation of state. The equation of state is normally modeled as a linear relation  $p = w\rho$ , where the parameter  $w$  takes the values 0, 1/3 and  $-1$  for the pressureless matter (applicable to both baryonic matter and CDM), for the radiation, and for the vacuum energy, respectively. Assuming the above mentioned separate components with the corresponding equations of state

---

<sup>2</sup>In fact, the Einstein field equations do not allow time-independent homogeneous and isotropic solutions except for the trivial case of a completely empty Universe.





**Figure 1.1.:** “Cosmic triangle” from the 1999 Science magazine paper by Bahcall et al. (1999). Three independent sources of information: the CMB fluctuations, Type Ia supernovae, and the abundance of the massive galaxy clusters seem to point towards the low matter density, spatially flat cosmological model. Spatial flatness is achieved by adding to the  $\sim \frac{1}{3}\rho_c$  of matter  $\sim \frac{2}{3}\rho_c$  contribution from the vacuum energy.

the most useful form for the Friedmann equation can be written as:

$$H(z) \equiv \frac{\dot{R}}{R} = H_0 E(z), \quad (1.3)$$

where

$$E(z) \equiv \sqrt{\Omega_r(1+z)^4 + \Omega_m(1+z)^3 + \Omega_k(1+z)^2 + \Omega_{DE} \exp \int_0^{\ln(1+z)} 3[1+w_{DE}(x)] dx}. \quad (1.4)$$

Here  $H(z)$  is the Hubble parameter describing the expansion rate,  $z = \frac{R}{R_0} - 1$  is the redshift, and  $\Omega_i$  is used to denote the density of the  $i$ -th component (r–radiation, m–matter, DE–Dark Energy) in units of the critical density  $\rho_c = \frac{3H_0^2}{8\pi G}$ . The most favorable value for the  $z = 0$  Hubble parameter  $H_0$  is found to be around 70 km/s/Mpc (e.g. Freedman et al. 2001, Spergel et al. 2003). The quantity  $\Omega_k$  is defined as:

$$\Omega_k \equiv 1 - \Omega_r - \Omega_m - \Omega_{DE}. \quad (1.5)$$

In terms of  $\Omega_k$  the curvature radius at  $z = 0$ ,  $R_0$ , can be expressed as:

$$R_0 = \frac{d_H}{\sqrt{|\Omega_k|}}, \quad (1.6)$$

where the Hubble distance

$$d_H = \frac{c}{H_0} = 2997.9 h^{-1} \text{ Mpc}. \quad (1.7)$$

$H_0$  is usually expressed as  $H_0 = h \cdot 100 \text{ km/s/Mpc}$ . The CDM and baryons both are modeled as pressureless components, i.e.  $\Omega_m = \Omega_{CDM} + \Omega_b$ , and the radiation component,  $\Omega_r$ , is the sum of photons and neutrinos:  $\Omega_r = \Omega_\gamma + \Omega_\nu$ . In Eq. (1.4) we have written the term for the DE in a very general form that allows for an arbitrarily varying equation of state parameter  $w_{DE}(z)$ . Since the currently available cosmological data is not yet able to provide a reasonable measurement of the function  $w_{DE}(z)$ , we assume a constant effective equation of state parameter  $w_{\text{eff}}$  throughout this thesis, instead.<sup>3</sup> In this case the last term in Eq. (1.4) reduces to  $\Omega_{DE}(1+z)^{3(1+w_{\text{eff}})}$ , which is just a constant  $\Omega_\Lambda$  in the simplest case of the  $\Lambda$ CDM model (since then  $w_{DE} \equiv -1$ ).

The analysis of the CMB temperature fluctuations measured by the WMAP satellite in combination with various other cosmological sources (e.g Type Ia supernovae, galaxy redshift surveys) has lead to the precise measurement of several density parameters. Currently most favored values read as:  $\Omega_m = 0.27 \pm 0.04$ ,  $\Omega_b = 0.044 \pm 0.004$ ,  $\Omega_\Lambda = 0.73 \pm 0.04$ ,  $\Omega_k = 0.02 \pm 0.02$  (Spergel et al. 2003). Including also the precise measurement of the dimensionless Hubble parameter  $h = 0.71 \pm 0.04$  (Spergel et al. 2003), the resulting smooth background model is able to accommodate the majority of the cosmological observations. That the spatially flat, low matter density model is able to perform remarkably well was already known at the end of 90's. This is demonstrated in Fig. 1.1 which is taken from the famous paper by Bahcall et al. (1999). Thus the rough “recipe” for the Universe is as follows: take  $\sim \frac{2}{3}\rho_c$  of vacuum energy, add  $\sim \frac{1}{3}\rho_c$  of CDM, and finally, as a spice, add a little bit of baryons.

In fact, the best determined density parameter is the one corresponding to the photons, i.e  $\Omega_\gamma$ . Since the measured CMB spectrum is practically a perfect black body with a temperature  $T_0 = 2.725 \pm 0.001 \text{ K}$  (Fixsen & Mather 2002), one can immediately calculate the physical energy density provided by the photons:  $\rho_{\gamma 0} = (4.642 \pm 0.007) \times 10^{-34} \text{ g/cm}^3$ , and thus  $\Omega_\gamma = (2.471 \pm 0.004) \times 10^{-5} h^{-2}$ . In the standard model the energy density in neutrinos is 68% of that in photons, and so  $\Omega_r = 1.68 \cdot \Omega_\gamma$ . Thus the contribution of the radiation to the total energy density of the current Universe is negligible. As seen from Eq. (1.4) this is not the case at higher redshifts: at redshifts  $z \approx 3300$  radiation starts to dominate.

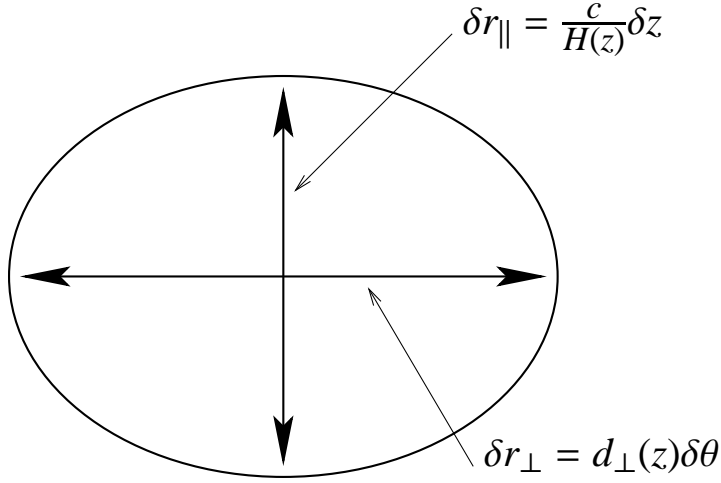
Using Eqs. (1.1), (1.2) and (1.3) one can express the comoving distance along the line of sight as:

$$d_{\parallel}(z) = d_H \int_0^z \frac{dz'}{E(z')}, \quad (1.8)$$

where  $d_H$  is the Hubble distance as defined above. The angle  $\delta\theta$  subtended by the comoving distance element with length  $\delta r_{\perp}$  perpendicular to the line of sight at redshift  $z$  can be given as:

---

<sup>3</sup>Throughout this thesis we also use the notation  $w_0$  in place of  $w_{\text{eff}}$ . The subscript <sub>0</sub> here refers to the lowest order term in some form of the series expansion of the function  $w_{DE}(z)$ .



**Figure 1.2.:** The relation of the comoving interval along,  $\delta r_{\parallel}$ , and perpendicular to the line of sight,  $\delta r_{\perp}$ , to the various kinematic measures of the expanding Universe.

$\delta \theta = \frac{\delta r_{\perp}}{d_{\perp}(z)}$  (see Fig. 1.2), where  $d_{\perp}$  is called the comoving angular diameter distance. Using the expression for the curvature radius in Eq. (1.6), one can write for  $d_{\perp}$ :

$$d_{\perp}(z) = R_0 S_k \left( \frac{d_{\parallel}(z)}{R_0} \right), \quad (1.9)$$

which in the case of the flat spatial sections reduces to  $d_{\perp}(z) = d_{\parallel}(z)$ , as expected. There are several classical cosmological tests that attempt to constrain the expansion law by directly measuring the redshift-distance relation. For this purpose one needs objects whose intrinsic luminosity or size is known. The first class of objects are called “standard candles”, while the second are known as “standard rulers”. The currently best known objects to qualify as standard candles are Type Ia supernovae.<sup>4</sup> Having standard candles available, one can measure directly the redshift dependence of the luminosity distance:

$$d_L(z) = (1 + z) d_{\perp}(z). \quad (1.10)$$

Below we show that a very good standard ruler is provided by the sound horizon, i.e. the distance a sound wave can travel since the Big Bang. Knowing the projected comoving length of the ruler perpendicular to the line of sight,  $\delta r_{\perp}$ , one can similarly measure the redshift dependence of the comoving angular diameter distance as given in Eq. (1.9) (cf. Fig. 1.2). Knowing the comoving extent of the ruler along the line of sight,  $\delta r_{\parallel}$ , and determining observationally the corresponding redshift interval  $\delta z$ , one can immediately find the Hubble parameter at the redshift of the ruler:  $H(z) = \frac{c \delta z}{\delta r_{\parallel}}$  (cf. Fig. 1.2). It is worth pointing out that  $H(z)$  is much more sensitive to  $w_{DE}(z)$  than  $\delta r_{\perp}(z)$ , since it involves only a single integration over the redshift, as can be seen from Eqs.

<sup>4</sup>To be precise, they are only “standardizable candles”, since the intrinsic luminosity for each object can only be estimated after applying the empirically determined relation between the peak brightness and the decline rate of the optical light curve.

(1.3) and (1.4), instead of two integrals, as in Eqs. (1.9), (1.8), and (1.4). If instead of the “real” standard ruler we have available an “object” (e.g. two point correlation function of galaxies) whose comoving shape is known, e.g.  $\delta r_{\parallel}/\delta r_{\perp} = 1$ , we can only find the combination  $H(z)d_{\perp}(z)$ . This is known as the Alcock-Paczynski test (Alcock & Paczynski 1979). Thus in the case of the “real” standard ruler we can perform an “absolute” version of the Alcock-Paczynski test, i.e. can get both  $H(z)$  and  $d_{\perp}(z)$  separately. Yet another classical test is based on a number count of objects. Namely if we have some class of objects whose comoving number density with its possible evolution is well known, then the observed number counts as a function of redshift gives us a direct information about the evolution of the comoving volume element with time.

Distance/Hubble parameter/volume element measurements at a fixed redshift will constrain only a single linear combination of the cosmological parameters. In order to determine various parameters separately one has to perform measurements at various redshifts (especially important is to probe higher redshifts that are out of the simple Hubble flow) and/or include information from other cosmological sources.

In the above discussion we assumed the validity of the General Relativity, which itself is not well tested at cosmologically large scales. This led us to the Friedmann equation (see Eqs. (1.3), (1.4)). One might take instead a different approach and assume only the validity of homogeneity and isotropy together with the assumption that space-time is described by a metric theory. This allows us to introduce the metric as in Eq. (1.1). Taking this kinematical approach and expanding the free function  $R(t)$  as follows:

$$R(t) = R_0 \left[ 1 + H_0(t - t_0) - \frac{1}{2}q_0 H_0^2(t - t_0)^2 + \frac{1}{6}j_0 H_0^3(t - t_0)^3 + \dots \right], \quad (1.11)$$

we can write the analog of Eqs. (1.3) and (1.4) in the form:

$$H(z) = H_0 \left[ 1 + (1 + q_0)z + \frac{1}{2}(j_0 - q_0^2)z^2 + \dots \right]. \quad (1.12)$$

The quantity  $q_0 \equiv -\frac{\ddot{R}}{RH^2}(z=0)$  is known as the deceleration parameter and  $j_0 \equiv -\frac{\dddot{R}}{RH^3}(z=0)$  the so-called jerk (Blandford et al. 2005) at the current epoch. If the observations are made at higher redshift, one certainly would not expand around  $z=0$ , but around e.g. the median redshift, instead. Expanding Eqs. (1.3) and (1.4) also up to the 2nd order in  $z$  and assuming constant effective equation of state parameter  $w_{\text{eff}}$ , we can find the correspondences:

$$q_0 = \frac{1}{2}\Omega_m + \frac{1+3w_{\text{eff}}}{2}\Omega_{DE}, \quad (1.13)$$

$$j_0 = \Omega_m + \left[ 1 + \frac{9}{2}w_{\text{eff}}(w_{\text{eff}} + 1) \right] \Omega_{DE}. \quad (1.14)$$

Throughout most of this thesis we assume flat spatial sections, i.e.  $\Omega_m + \Omega_{DE} = 1$ , and constant effective equation of state parameter  $w_{\text{eff}}$ . Under these assumptions one can easily replace the kinematic parametrization  $(h, q_0, j_0)$  with the dynamic one  $(h, \Omega_m, w_{\text{eff}})$ . The inverse relations to Eqs. (1.13) and (1.14) can then be given as:

$$\Omega_m = \frac{2[j_0 - q_0(1 + 2q_0)]}{1 + 2(j_0 - 3q_0)}, \quad (1.15)$$

$$w_{\text{eff}} = \frac{2(3q_0 - j_0) - 1}{3(1 - 2q_0)}. \quad (1.16)$$

If  $w_{\text{eff}}$  is restricted to the range  $w_{\text{eff,min}} \leq w_{\text{eff}} \leq 0$ , then in the currently considered case  $\frac{3w_{\text{eff,min}}+1}{2} \leq q_0 \leq \frac{1}{2}$  and  $-\frac{1}{8} \leq j_0 \leq (3w_{\text{eff,min}}+1)(\frac{3w_{\text{eff,min}}}{2}+1)$ . The observationally favored spatially flat  $\Lambda$ CDM model corresponds to  $q_0 = \frac{3}{2}\Omega_m - 1$  and  $j_0 = 1$ .

The above described homogeneous and isotropic cosmological models are known as Friedmann-Lemaître-Robertson-Walker (FLRW) models.

### 1.1.2. Perturbed FLRW models

FLRW models form the cornerstone of the standard Big Bang theory. Big Bang theory, although being very successful in predicting correctly e.g. the abundance of light elements and the existence of the CMB, has several problems it is not able to address properly. For example: why are the spatial sections of our Universe so close to being flat? Under the usual Einstein-de Sitter (i.e.  $\Omega_{\text{tot}} \simeq \Omega_m \simeq 1$ ) behavior of the early Universe tiny deviations from the spatial flatness would be amplified very quickly, making our well studied late-time Universe very unnatural outcome. This problem is known as the flatness problem. The second unanswered question is why our Universe appears so homogeneous at large scales, although under the decelerated expansion of the early Universe these large regions had no chance of exchanging information. This is known as the horizon problem. There are also several other problems and it turns out that many of these can be naturally solved by postulating an early period of accelerated expansion. The models with an early accelerated expansion phase are known as the inflationary models. Here we do not delve into the vast subject of inflationary cosmology, but refer the reader to the standard texts like [Linde \(1990\)](#), [Kolb & Turner \(1990\)](#), [Liddle & Lyth \(2000\)](#), instead. The only important point for us to stress is that inflation provides a causal mechanism for generating initial density fluctuations (and thus providing the solution to the so-called fluctuation problem) that under the gravitational instability develop into the large-scale structure that we observe in the present Universe. As the inflation is stretching the tiny quantum fluctuations to the vast cosmological scales, we might really have a good chance of learning something about these exotic high energy processes taking place in the early Universe through the study of the large-scale structure, or the CMB angular fluctuations. The simplest inflationary scenarios predict adiabatic and Gaussian initial conditions with a nearly scale-invariant Harrison-Zeldovich spectrum (i.e. the spectral index  $n_s = 1$ ) ([Harrison 1970](#), [Zeldovich 1972](#)). All of these predictions are currently completely compatible with the available observational constraints. Also the generic prediction for the nearly flat spatial sections is in full agreement with the observational data. Thus the observations indeed seem to require an early accelerated epoch of the Universe. The question is of course how much we can possibly learn about the extremely high energy physical processes that are involved, and that in the case when almost all the characteristic observational features are very close to their simplest “vanilla” values.

As the initial fluctuations are given by the Gaussian random field with zero mean, they are completely described by the two-point function, which in the Fourier space is known as the power spectrum. The Fourier space picture is favorable since the covariance matrix of the Fourier modes has only diagonal elements. Moreover, this covariance structure is preserved under linear evolution since each of the Fourier modes (as being eigenmodes of the flat space Laplacian) evolves independently. The initial spectrum is usually parametrized as:

$$P_{\text{init}}(k) = A_s k^{n_s}, \quad (1.17)$$

where the spectral index  $n_s = 1$  corresponds to the scale invariant (Harrison-Zeldovich) case. Here we consider only scalar perturbations since these are the ones relevant for the studies of the large-scale structure. In addition to these inflation also excites tensor modes, i.e. gravitational waves, and there is a hope in the “CMB community” to detect these through the CMB  $B$ -mode (i.e. divergence-free) polarization by the extremely sensitive future experiments like CMBPol<sup>5</sup>. Inflation does not excite vector modes, and even if there are some to start with, these should decay due to the expansion of the Universe, unless they are regenerated by some form of “active perturbations” that can be provided by cosmic defect models. The defect models are strongly disfavored as the dominant source for the initial fluctuations by the CMB measurements (e.g. [Durrer et al. 2002](#)).

In order to calculate how these initial perturbations evolve in time, one has to integrate the coupled set of Einstein-Boltzmann equations. Here one has to specify all the separate matter/energy components with possible interactions between them. The usual treatment includes baryons, photons, CDM, and neutrinos. All the electrically charged species in the radiation dominated plasma are strongly coupled by Coulomb interaction. The other important process is Compton scattering of electrons and photons that together with Coulomb interaction keeps photons and baryons tightly coupled up to redshifts  $\sim 1200$ . When the temperature of the Universe drops below  $\sim 0.3$  eV, hydrogen starts to recombine, eliminating the population of free electrons. This leads to the break-down of the tightly coupled baryon-photon fluid and to the enormous increase of the mean free path of the photons. These freely streaming photons are the ones we observe as the CMB. The baryons that cease to feel the strong pressure of the photon gas start to fall back onto the CDM density peaks that were able to grow significantly since the matter-radiation equality. The CDM component is strongly needed to reconcile the small fluctuations in the CMB with the highly evolved large-scale structure that surrounds us. With baryons only one cannot obtain fast enough growth. This is actually one of the strongest arguments that requires the existence of the component with the properties of the CDM.

The detailed treatment of the evolution of the initial perturbations is a fairly technical subject that is beyond the scope of the current short introduction. There exist several excellent sources that present the perturbation calculation in full glory e.g. [Mukhanov et al. \(1992\)](#), [Ma & Bertschinger \(1995\)](#). We are mostly concerned with the evolution of the matter density fluctuations. The results of the full calculation can be expressed as a transfer function  $T(k)$  that maps the initial matter fluctuation spectrum to the linearly evolved spectrum at the current epoch. Since well inside the regime of matter domination the shape of the power spectrum does not change anymore one can express the low redshift power spectrum as:

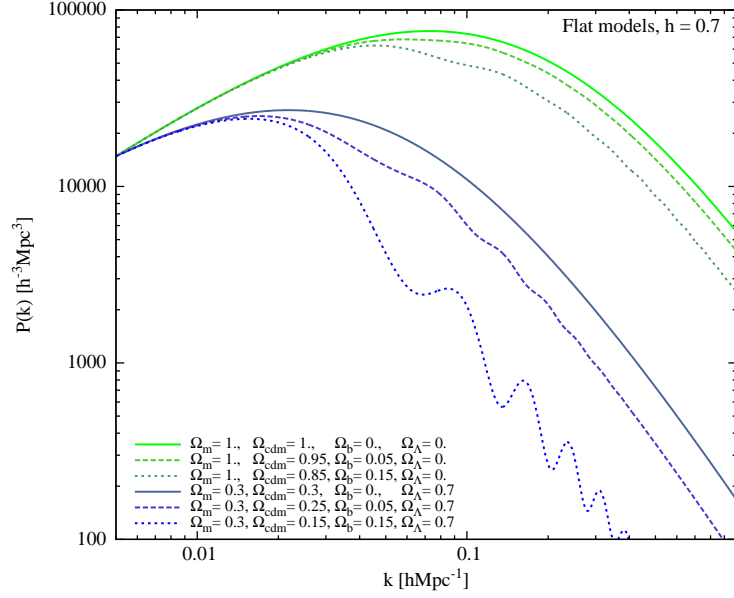
$$P(k, z) = \frac{g^2(z)}{g^2(0)} T^2(k) P_{\text{init}}(k), \quad (1.18)$$

where the initial power spectrum is usually taken in the form of Eq. (1.17), and the linear growth factor  $g(z)$  satisfies the following differential equation:

$$\frac{d^2 g}{d \ln a^2} + \left\{ 1 - \frac{1}{2} [\Omega_m(z) + (1 + 3w(z)) \Omega_{DE}(z)] \right\} \frac{dg}{d \ln a} - \frac{3}{2} \Omega_m(z) g, \quad (1.19)$$

---

<sup>5</sup><http://universe.nasa.gov/program/inflation.html>



**Figure 1.3.:** Various linear matter power spectra for the spatially flat models with the Harrison-Zeldovich initial spectrum and with the dimensionless Hubble parameter  $h = 0.7$ , normalized to the same amplitude at the largest scales. The upper group of curves corresponds to the Einstein-de Sitter (i.e.  $\Omega_m = 1$ ), while the lower group to the  $\Lambda$ CDM models with  $\Omega_m = 0.3$  and  $\Omega_\Lambda = 0.7$ . Within each group the amount of baryons increases from the top to the bottommost curve.

where  $a = \frac{1}{1+z}$ , and  $\Omega_m(z)/\Omega_{DE}(z)$  is the matter/dark energy density parameter at redshift  $z$ , i.e.:

$$\Omega_m(z) = \Omega_m \frac{(1+z)^3}{E^2(z)}, \quad (1.20)$$

$$\Omega_{DE}(z) = \Omega_{DE} \frac{\exp \int_0^{\ln(1+z)} 3 [1 + w_{DE}(x)] dx}{E^2(z)}. \quad (1.21)$$

In case the dark energy is provided by the cosmological constant, the solution for Eq. (1.19) can be given as:

$$g(a) = \frac{5}{2} \Omega_m H_0^2 H(a) \int_0^a \frac{da'}{(a' H(a'))^3}. \quad (1.22)$$

Here the growth function is normalized such that at early times (i.e. well in the Einstein-de Sitter phase)  $g(a) = a$ .

One can use several publicly available Boltzmann codes like COSMICS<sup>6</sup> (Bertschinger 1995) and the newer generation of tools partly based on it, such as CMBFAST<sup>7</sup> (Seljak & Zaldarriaga

<sup>6</sup><http://web.mit.edu/edbert/>

<sup>7</sup><http://www.cmbfast.org/>



1996), CAMB<sup>8</sup> (Lewis et al. 2000), CMBEASY<sup>9</sup> (Doran 2005), to calculate transfer functions. As the names of many of these tools hint they can also be used for calculating the CMB angular power spectra for various cosmological models. While the older solvers for the CMB angular spectra used to integrate the full Boltzmann hierarchy up to the desired multipole  $\ell$ , the newer tools separate the problem into the source term and the subsequent radiative transfer towards the low redshift observer (Seljak & Zaldarriaga 1996). There also exist several approximate and very fast tools for the CMB angular spectrum calculation such as CMBfit<sup>10</sup> (Sandvik et al. 2004), DASH<sup>11</sup> (Kaplighat et al. 2002) and CMBwarp<sup>12</sup> (Jimenez et al. 2004). Similarly, there exist accurate fitting formulae for the matter transfer functions by e.g. Eisenstein & Hu (1998) and Novosyadlyj et al. (1999). All of these approximate tools are useful for the fast likelihood calculations in the high dimensional cosmological parameter spaces. Many of the approximations are based on various analytical results (see eg. Hu 1995, Hu & Sugiyama 1995, 1996, Eisenstein & Hu 1998) that significantly help to improve our knowledge about the underlying physical processes. In the current thesis we make extensive use of the CAMB code to calculate the CMB angular power spectra and the matter transfer functions.

A few example power spectra for the CDM models are shown in Fig. 1.3. Here all the models are spatially flat with the Harrison-Zeldovich initial spectrum and the dimensionless Hubble parameter  $h$  is fixed to 0.7. The upper group of curves corresponds to the Einstein-de Sitter models (i.e.  $\Omega_m = 1$ ) whereas the lower to the  $\Lambda$ CDM models. Inside each group the amount of matter provided by the baryons is increasing from the top to the bottommost curve. All the spectra have been normalized to the same amplitude at the largest scales. Some important things to notice: (i) the power spectra for the high matter density models turn over at smaller scales (i.e. higher  $k$ ) than the spectra for the lower density ones; (ii) the higher baryon fraction  $\Omega_b/\Omega_m$  leads to the stronger damping of the spectrum; (iii) the models with the high baryon fraction also develop a periodic sequence of acoustic oscillations. The turn-over in the matter spectrum occurs at the scale corresponding to the size of the horizon at matter-radiation equality, which itself is inversely proportional to the parameter combination  $\Gamma \equiv \Omega_m h$  (assuming distances are measured in units of  $h^{-1}$  Mpc), the so-called shape parameter. The reason for the power spectrum to turn over is the fact that during radiation domination the fluctuation modes inside the horizon, due to the high pressure of the photon gas, can grow only logarithmically, whereas the ones outside the horizon keep on growing proportional to the scale factor squared. After the matter-radiation equality the fluctuations in the CDM component can start growing more efficiently. As the baryons are still coupled to the photon gas the baryonic fluctuations cannot yet start growing. This continues until the epoch of recombination when the Universe turns neutral and baryons get released from the pressure of the photon gas, allowing them to start falling into the CDM potential wells. Since in the “concordance” cosmological model baryons make up  $\sim 15\%$  of the total matter, they also have some non-negligible gravitational influence on the CDM component, leading to some “smoothing” of the CDM density field. The more baryons, the stronger the smoothing/damping of the final matter spectrum. The formation of the acoustic features in the

---

<sup>8</sup><http://camb.info/>

<sup>9</sup><http://www.cmbeasy.org/>

<sup>10</sup><http://www.hep.upenn.edu/sandvik/CMBfit.html>

<sup>11</sup><http://bubba.ucdavis.edu/DASH/>

<sup>12</sup><http://www.physics.upenn.edu/raulj/CMBwarp/>

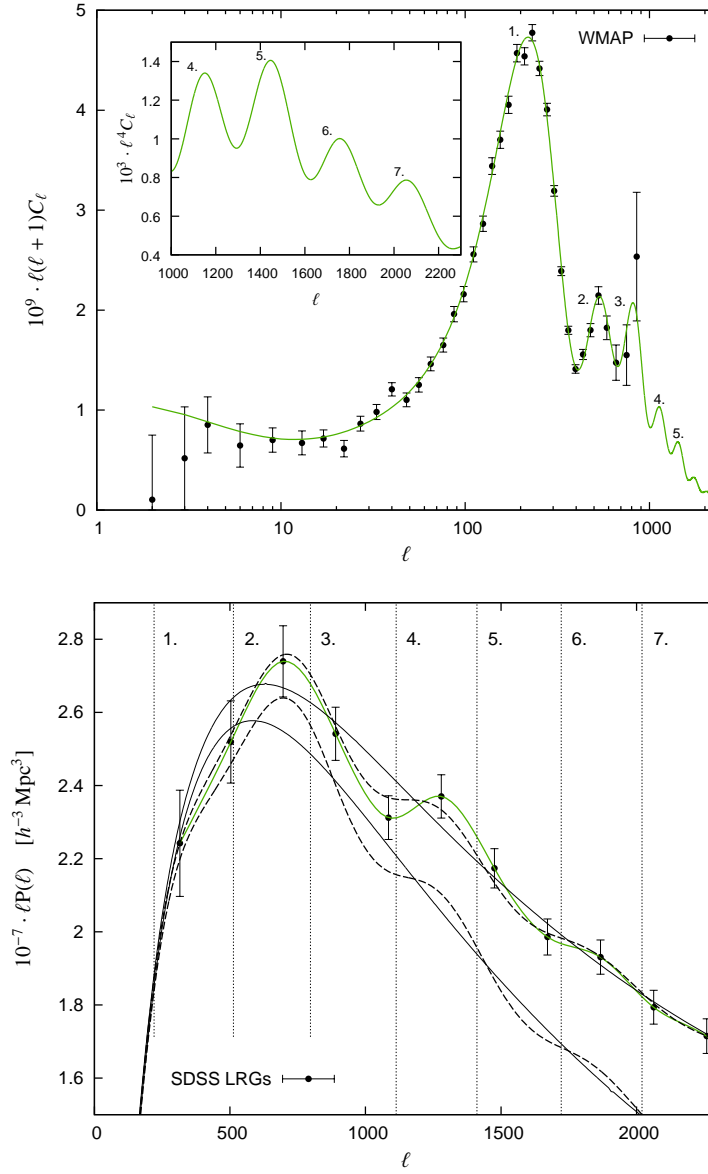


matter power spectrum will be discussed in some detail in the next section.

## 1.2. Introduction to acoustic oscillations

During the last decade observational cosmology has witnessed extremely rapid development. Currently several key parameters describing our Universe, such as the global densities of various matter/energy components, Hubble parameter, etc. have been measured to an accuracy better than 10%. This rapid progress is largely driven by precision measurements of the angular temperature fluctuations of the Cosmic Microwave Background (CMB). Since the density inhomogeneities at the time when the Universe recombined and the CMB was emitted (at redshift  $z \sim 1100$ ) were only one part in 100,000 the Einstein-Boltzmann equations describing the evolution of these fluctuations can be linearized to a very good approximation, which makes the solution of this complicated equation set possible. In general, the success of the current theoretical models in explaining many of the observed features of the CMB to a high precision can be seen as a guarantee that these models are on the right track. We have learned several important things about the primordial seed fluctuations, presumably generated during the early accelerating expansion phase of the Universe, in analyzing these small fluctuations imprinted onto the CMB sky. First, the CMB data demands the dominant fluctuation mode to be adiabatic, i.e. the initial number-density fluctuations  $\frac{\delta n_i}{n_i}$  for all components of the cosmic fluid follow each other.<sup>13</sup> Second, the initial fluctuations are compatible with being Gaussian with a roughly scale-free power spectrum. These observations are in full agreement with the predictions of the simplest inflationary Universe models. As the seed fluctuations are of Gaussian nature they are fully described by the two-point function (since we are defining the fluctuations around the mean CMB temperature, the mean fluctuation itself is zero), which in harmonic space is known as the power spectrum. Roughly speaking, the power spectrum shows how the fluctuation power is distributed amongst the perturbations having different wavelengths. The angular spectrum of the CMB temperature fluctuations as measured by the WMAP team together with the best fitting  $\Lambda$ CDM model with approximately scale-invariant adiabatic initial conditions is shown in the upper panel of Fig. 1.4. Here the most characteristic features are the so-called acoustic peaks with the most prominent first peak corresponding to an angular scale of  $\sim 0.6^\circ$ . Thus the temperature fluctuations of the CMB sky have the strongest contrast (after removing the dipolar temperature anisotropy caused by the motion of the Local Group) for the patches with a typical size of  $\sim 0.6^\circ$ . This typical scale is directly related to the distance the sound waves in the tightly coupled baryon-photon fluid in the pre-recombination Universe can have traveled since the Big Bang. According to the

<sup>13</sup>This implies that fluctuations in the mass density  $\delta_i \equiv \frac{\delta \rho_i}{\rho_i}$  for photons, neutrinos, CDM, and baryons are related as  $\delta_\gamma = \delta_\nu = \frac{4}{3}\delta_{CDM} = \frac{4}{3}\delta_b$ . These fluctuations are directly related to the perturbations in the spatial curvature. There is also possibility for the initial fluctuation modes that do not perturb spatial curvature. These are known as isocurvature perturbations, and they can be seen as entropy perturbations  $S_i \equiv \frac{\delta n_i}{n_i} - \frac{\delta n_\gamma}{n_\gamma}$ . So for the  $N$ -component “cosmic soup” in addition to the adiabatic mode there can be up to  $N - 1$  different isocurvature modes. The most general perturbation can always be expressed as a combination of the adiabatic mode with several isocurvature components. All the currently existing observational data are consistent with the adiabatic initial conditions, as predicted by the simplest inflationary models. According to the CMB data the dominant adiabatic fluctuation mode can have only a small isocurvature admixture. For this reason we consider only models with pure adiabatic initial conditions throughout this thesis.



**Figure 1.4.:** Upper panel: The CMB angular power spectrum as measured by the WMAP team together with the best-fitting  $\Lambda$ CDM model curve. The inset shows the zoom into the damping tail. Due to very strong decline of the CMB angular spectrum at large multipoles the  $y$ -axis is multiplied with an extra factor of  $\ell^2$  in comparison to the main figure. Lower panel: The power spectrum of the SDSS LRGs plotted in a way allowing for a direct comparison with the corresponding CMB spectrum provided in the upper panel. The comoving wavenumber  $k$  was transformed to the multipole number  $\ell$  such that  $\ell \simeq 9940 \cdot k [h \text{Mpc}^{-1}]$ , where  $9940 h^{-1} \text{Mpc}$  is the comoving angular diameter distance to the last scattering surface for the best-fit WMAP “concordance” model (Spergel et al. 2003). The solid green line is the cubic spline fitted to the observational data. The lower dashed curve is the linearly evolved matter spectrum corresponding to the best fitting model from the panel above, while the upper dashed line shows the spectrum after incorporating the treatment for the redshift space distortions and nonlinear evolution. The thin solid lines represent the “smoothed” models without baryonic oscillations. All the model spectra here are convolved with a survey window function. The vertical dotted lines mark the positions of the acoustic peaks in the CMB power spectrum.

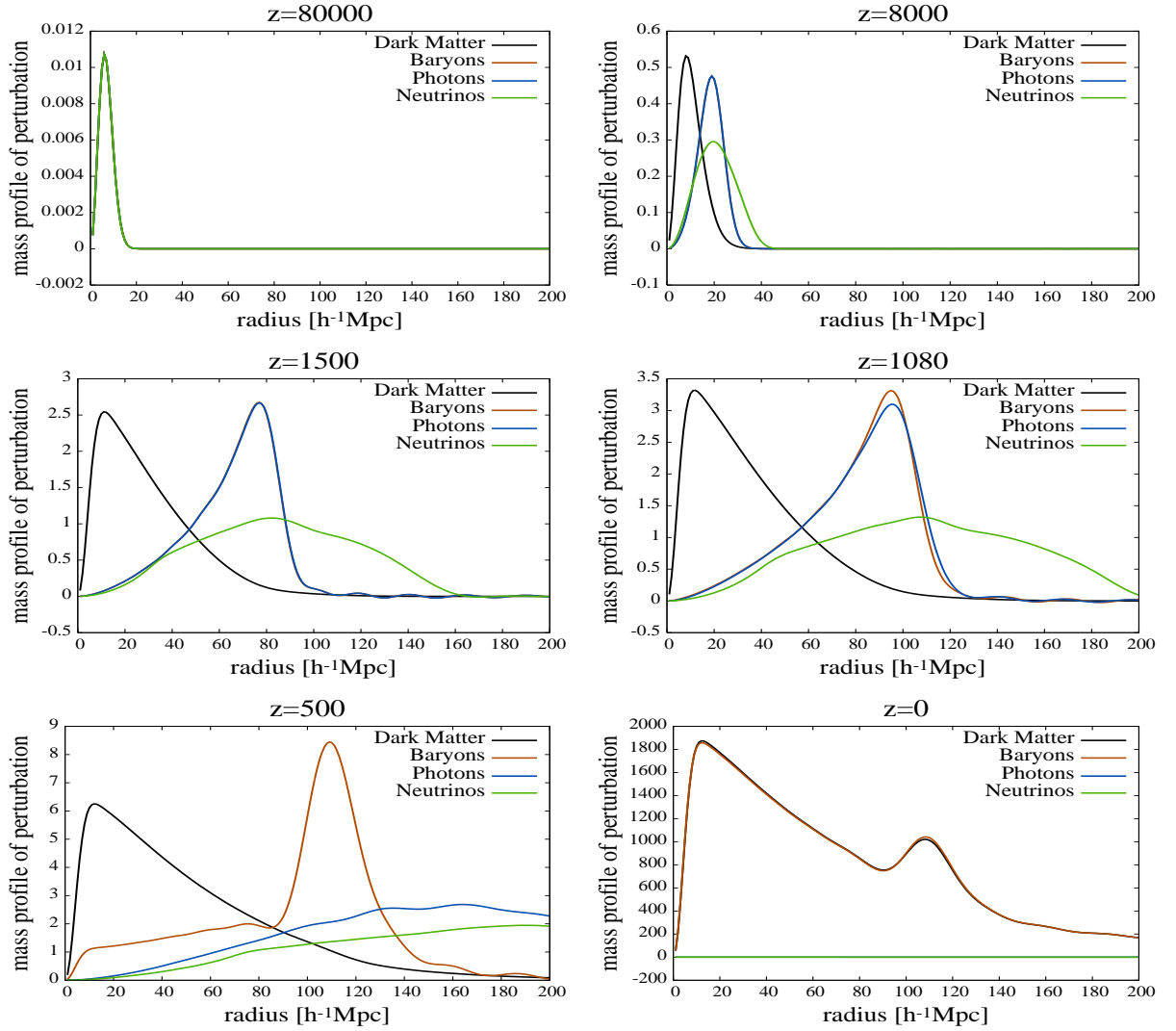
currently most favorable model for the large-scale structure formation– the gravitational instability theory– these tiny high-redshift fluctuations, as probed by the CMB, serve as seeds for the highly evolved cosmic structure surrounding us. The important consequence of the linear gravitational instability theory is that all the features present in the initial matter fluctuation spectrum should survive throughout cosmic evolution. The “concordance” cosmological model due to its relatively low baryonic matter fraction (only  $\sim 15\%$  of the matter is in the form of the baryons, the rest being contributed by the cold dark matter (CDM)) predicts that the low redshift matter power spectrum should contain small ( $\sim 5\%$ ) fluctuations due to the acoustic phenomena. And this is indeed the case, as has now been confirmed by analysis of the spatial clustering of the Sloan Digital Sky Survey (SDSS)<sup>14</sup> Luminous Red Galaxy (LRG) sample by Eisenstein et al. (2005) (who determined the two-point correlation function) and more recently by Hütsi (2005), Hütsi (2006a) (who determined both the power spectrum and two-point correlation function). Cole et al. (2005) have similarly detected acoustic oscillations in the power spectrum of the 2dF<sup>15</sup> galaxy sample. The SDSS LRG power spectrum as measured by Hütsi (2006a) is shown on the lower panel of Fig. 1.4. In order to ease the comparison with the upper panel we have converted the comoving wavenumbers to the corresponding multipole numbers, as explained in the figure caption. Here the lower dashed line corresponds to the linearly evolved matter power spectrum for the same  $\Lambda$ CDM model whose CMB power was shown on the upper panel. The upper dashed curve shows the model spectrum after the corrections for the nonlinear effects and redshift space distortions are taken into account. The solid green line represents the cubic spline fitted to the data points, vertical dotted lines give the locations of the CMB acoustic peaks, and the thin solid lines represent the “smoothed” models without baryonic oscillations. An important point to note here is the fact that the CMB spectrum has an oscillation frequency approximately two times higher than the corresponding frequency in the matter power spectrum. The reason for this will be explained below.

Usually the cosmological perturbation equations are solved in harmonic space, i.e. all the quantities are expressed as superpositions of plane waves (or their generalizations if the spatial sections of the space-time are not flat). This representation is very convenient for numerical studies, since due to the motion invariance of the evolution equations the time dependence factorizes out, and thus we are basically left with a standing wave decomposition whose amplitudes can just be appropriately adjusted as the time goes by. Although numerically convenient this representation is not very intuitive. A more enlightening picture arises if the analysis is carried out in real space instead. In Fig. 1.5 we show a pedagogical example by D.Eisenstein<sup>16</sup>, that itself is based on an original work by Bashinsky & Bertschinger (2002). The figure displays an evolution sequence of the initial adiabatic spherical density perturbation. Here the  $x$ -axis displays the comoving radius  $r$  and the  $y$ -axis presents the mass profiles of several perturbation components (as listed in the legend), i.e. the fractional overdensities  $\delta\rho_i/\rho_i$  times  $r^2$ , in arbitrary units. For clarity the extra factors of  $4/3$  have been omitted for the relativistic components. Initially all the components are confined inside the perturbation. Since the initial overdensity in the tightly coupled baryon-photon fluid also corresponds to the initial overpressure, an outward-moving spherical sound wave will be launched. As the neutrinos are not coupled to the rest of the matter at these

<sup>14</sup><http://www.sdss.org/>

<sup>15</sup><http://www.mso.anu.edu.au/2dFGRS/>

<sup>16</sup><http://cmb.as.arizona.edu/~eisenste/acousticpeak/>



**Figure 1.5.:** The evolution sequence of the initial adiabatic density peak. The x-axis shows the comoving distance in  $h^{-1}$  Mpc, while the y-axis displays the mass profile of the perturbation in arbitrary units. Black, red, blue and green lines correspond to the CDM, baryons, photons and neutrinos, respectively. The redshifts corresponding to each of the “snapshots” are given above each panel. The linearized Einstein-Boltzmann equation set was solved using publicly available CAMB software. (Example originally due to D.Eisenstein)

relatively low redshifts, they start to diffuse out of the sound wave. CDM distribution is also significantly “smoothed” due to the gravitational pull of the outward-traveling baryon-photon perturbation. This “smoothing” gives the final CDM spectrum with its characteristic turnover. At redshifts around  $z \sim 1050$  the tight coupling between the baryons and photons breaks down as the Universe starts to recombine and thus the photons can begin to diffuse out of the sound wave. As the baryons are released from the photon pressure the sound speed drops radically and the sound wave practically stalls. In fact the slow motion of the baryonic shell continues down to  $z \sim 200$  (the so-called drag epoch). After that redshift the baryons decouple from the photons completely and can start to fall back onto the central CDM density peak that has grown significantly since the matter-radiation equality. Because in the “concordance” cosmological model the baryonic density is not completely negligible it has also some gravitational effect on the CDM component. Thus the final density profile will have a small density enhancement at the distance corresponding to the size of the sound horizon at the end of the drag epoch. A more general initial density field can always be expressed as a superposition of  $\delta$ -spikes. As we are dealing with linear perturbation theory the evolved field can be expressed as the sum of the separately evolved  $\delta$ -functions, i.e. the ordinary Green’s function method.<sup>17</sup> Thus in the more general case one would also expect an enhancement in the two-point correlation function at the separation corresponding to the sound horizon. This relatively narrow peak in the correlation function leads to the oscillating behavior of the power spectrum shown in Fig. 1.4. The CMB sky in this picture corresponds to the superposition of the “photon shells” (whose thickness is determined by the efficiency of the diffusive processes) that are cut by the last scattering surface which itself is a shell with a thickness of  $\sim 30 h^{-1}$  Mpc. Since in the case of the “photon shells” no perturbation is left in the center the corresponding correlation length is approximately twice as large as the one for the matter component, and thus the CMB angular spectrum also fluctuates twice as frequently. To be more precise, the acoustic horizon as measured from the CMB sky is slightly smaller than the one imprinted in the matter distribution, since as mentioned above, the sound wave does not stall completely at recombination. Useful fitting formulae for the acoustic scales relevant for the CMB and large-scale structure studies can be found in Appendix A. There we also show how the directly observable quantities depend on various cosmological parameters.

The acoustic scale measured from the SDSS LRG power spectrum shown in Fig. 1.4 was found to be  $(105.4 \pm 2.3) h^{-1}$  Mpc. This value assumes that the background model used to calculate the distances is the WMAP best-fit “concordance” model. For different background models this value can easily be rescaled. It is worth pointing out that WMAP data<sup>18</sup> (Hinshaw et al. 2003, Kogut et al. 2003) together with a prior on the Hubble parameter from the HST Key Project<sup>19</sup>,  $H_0 = 72 \pm 8$  km/s/Mpc (Freedman et al. 2001), would predict the corresponding scale to be  $(107 \pm 20) h^{-1}$  Mpc. Thus the measurement given above provides a factor of  $\sim 10$  improvement. Also it turns out that the models with baryonic features are favored by  $3.3\sigma$  over their “smoothed-out” counterparts without any oscillatory behavior, i.e. the acoustic features are

<sup>17</sup>The curves in Fig. 1.5 are only approximations for the exact Green’s functions, since due to numerical convenience, as we use the spectral methods for solving the equations, our initial density peak was taken to be a Gaussian with a finite width. Thus these functions correspond to the Green’s functions convolved with a narrow Gaussian window, and as such, the late-time curves, because of being significantly broader than the initial Gaussian, provide already very good approximations.

<sup>18</sup><http://lambda.gsfc.nasa.gov/>

<sup>19</sup><http://www.ipac.caltech.edu/H0kp/>

**Table 1.1.:** The list of the most common cosmological parameters.

Inflationary parameters	$A_s, n_s, \frac{d \ln n_s}{d \ln k}, A_t, n_t$
Evolutionary parameters	$h, \Omega_m, \Omega_b, \Omega_r, \Omega_\nu, \Omega_{DE}, w_{\text{eff}}, \tau$

detected at a relatively high confidence level. All of this demonstrates the great promise of the future extremely large galaxy redshift surveys to be carried out with instruments like WFMOS (Wide-Field Multi-Object Spectrograph), formerly known as K.A.O.S. (Kilo-Aperture Optical Spectrograph) <sup>20</sup>.

The full consequences of the acoustic scale measurement, as quoted above, for the cosmological parameters will be worked out in the last chapter of this thesis.

### 1.3. Cosmological parameters. Markov Chain Monte Carlo

A few decades ago observational cosmology was termed to be a “quest for two parameters” only: the Hubble parameter  $H_0$  and the deceleration parameter  $q_0$  (see Eq. (1.11)). During the years our knowledge about the Universe has increased dramatically. Most importantly, we now seem to have a very good theory available to explain the growth of the tiny initial fluctuations, as observed at the last scattering surface of the CMB, to the large-scale cosmic structure surrounding us at low redshifts. We are now asking significantly more complex questions, and as such, also the parameter space has grown from two dimensions up to more than ten dimensional. The list of the most common parameters is given in Table 1.1. Here we have separated out the parameters that describe the initial fluctuations and called them “inflationary parameters”. These include the spectral indices ( $n_s, n_t$ ) and amplitudes ( $A_s, A_t$ ) of the scalar/tensor perturbations and the running of the scalar spectral index  $\frac{d \ln n_s}{d \ln k}$ . The other group of parameters determine the expansion law of the background Universe or/and the way perturbations evolve. We call these parameters “evolutionary parameters”. Here the list includes a dimensionless Hubble parameter  $h$ ; various density parameters  $\Omega_m, \Omega_b, \Omega_r, \Omega_\nu, \Omega_{DE}$  for the total nonrelativistic matter, baryons, radiation, massive neutrinos, and dark energy, respectively; dark energy equation of state parameter  $w_{\text{eff}}$ , which in the most general case can be a free function of redshift; and  $\tau$ , the optical depth to the last scattering surface, which influences the level of the CMB fluctuations on smaller scales.

In order to carry out parameter analysis, using e.g. CMB or galaxy clustering data, it is essential to have accurate and fast tools available for the calculation of the evolution history of the initial perturbations. A few of the publicly available tools were listed in subsection 1.1.2. In this thesis we almost always use the CAMB software package developed by Anthony Lewis <sup>21</sup> for this purpose. For the CAMB it takes less than a second on a 3 GHz processor to calculate the CMB angular power spectrum for the flat models. Including the calculation of the transfer functions, and/or calculating spectra for the non-flat models or models including massive neutrinos, takes somewhat longer time (in any case less than 10 seconds). It is now immediately clear that usual

<sup>20</sup><http://www.noao.edu/kaos/>

<sup>21</sup><http://camb.info/>



grid-based likelihood calculation is not feasible when using accurate Boltzmann solvers such as CAMB. If we have a 10-dimensional parameter space and in case we would calculate only 10 points per dimension, the total number of points would be  $10^{10}$ . Having the above 3 GHz serial machine we would have to wait  $\sim 300$  years to finish the calculation. Also the storage of the resulting grid would require  $\sim 40$  TB of disk space. This task would be barely within the reach of the current day biggest supercomputers, but fortunately there exist much better numerical techniques that require orders of magnitude less computational power.

Almost all the recent cosmological parameter studies use the Markov Chain Monte Carlo (MCMC) techniques, well known in the field of computational statistical mechanics, to speed up the likelihood calculations. For a good overview of the MCMC methods see e.g. Neal (1993), Gilks et al. (1996), MacKay (2003). Here instead of calculating the likelihood function at the predefined set of grid points one draws a sample of events from the posterior probability (posterior probability = likelihood times the prior probability) distribution. Having a sample of events from the posterior distribution makes the calculations of the marginal distributions, means, standard deviations, etc. completely trivial. The question is of course how to set up an algorithm that moves around in the high dimensional parameter space and as an output provides us with a fair sample from the posterior under investigation. There are several sampling algorithms available (see e.g. Neal 1993, 2000) with the most famous amongst these being the algorithm due to Metropolis and Hastings (MH) (Metropolis et al. 1953, Hastings 1970). The MH in its most basic form is extremely simple:

1. Start at some point  $\Theta_i$ <sup>22</sup> ( $i = 1$ ) in a high dimensional parameter space.
2. Draw a step  $\Delta\Theta$  from the proposal distribution  $p(\Delta\Theta)$ .
3. If the proposed new point  $\Theta_{i+1} = \Theta_i + \Delta\Theta$  has higher likelihood than  $\Theta_i$ , i.e.  $\mathcal{L}(\Theta_{i+1}) > \mathcal{L}(\Theta_i)$ , accept it as a new member of the chain. Otherwise accept it only with probability  $\frac{\mathcal{L}(\Theta_{i+1})}{\mathcal{L}(\Theta_i)}$ . If point gets rejected take  $\Theta_{i+1} = \Theta_i$ . The point is always rejected if it violates some prior constraints.
4. Go to step 2.

It can be shown that the asymptotic distribution for the Markov chain obtained this way agrees with the posterior distribution we are investigating. There are several technical issues: e.g. how to assure that the chain has reached an equilibrium? For the final analysis the initial transient period when the chain has not yet equilibrated– the so-called “burn in” period– is removed from the chain. There exist several “recipes” for estimating the length of the “burn in” (e.g. Raftery & Lewis 1995). It is evident that the neighboring elements of the chain are not independent. To remove these correlations one usually “thins the chain”, i.e. uses only every  $n$ –th element of the chain. Also there exist “recipes” to find appropriate thinning factor (e.g. Raftery & Lewis 1995). The hardest part of the implementation of the MCMC algorithm is to find an appropriate proposal distribution. The width of the proposal distribution must approximately match the width of the posterior we try to sample from. In case of too broad proposal distribution the chain gets stuck to one point for a long time, leading to a very strongly correlated chain with only a small number of independent samples. This is termed as an insufficient mixing of the chain. Insufficient mixing

---

<sup>22</sup>Here  $\Theta_i$  denotes the value of the parameter vector at step  $i$ .

occurs also in the opposite case of the too narrow proposal distribution. In that case, as almost all of the points will get accepted <sup>23</sup>, the chain performs a Brownian motion with a too small step size to sample properly the whole region where the posterior has significant amplitude. In the case of the cosmological parameter estimation one often uses a multidimensional Gaussian as a proposal distribution. As we often have a rather good knowledge about possible parameter degeneracies, it is useful to transform to the independent normal coordinates, which leads to a much better mixing. Without this transformation it is very hard to achieve a good mixing along degenerate directions. Since the topology of the “acceptance region” in the cosmological parameter space is usually relatively simple, it is completely acceptable to use the simple MH algorithm with the Gaussian proposal distribution. In the case of more complex topologies, e.g. several disconnected regions with significant posterior probability, one has to rely on more advanced sampling techniques (see e.g. Neal 1993, 2000)

Concerning cosmological parameter estimation, one often gets a rather good sampling of a 10–dimensional posterior with a chain of a few  $10^5$  elements. Compared to the above mentioned grid-based technique this leads to a huge increase in performance: now we have to calculate the likelihood in a factor of  $\sim 10^5$  times less number of points. In general, the number of points needed for the MCMC algorithm to sample the posterior distribution satisfactorily depends approximately linearly on the dimensionality of the parameter space. In contrast, for the usual grid-based methods this scaling is exponential. Moreover, quite often it is possible to “reuse” already calculated Markov chains. For example, if we obtain some additional data and the posterior of the complete dataset stays relatively close to the old one, we can simply reweight the available chain instead of building the new chain “from the scratch”. This method is known as Importance Sampling. For more details see e.g. Gilks et al. (1996). In this thesis we perform the MCMC calculations using the publicly available Cosmomc<sup>24</sup> package (Lewis & Bridle 2002) with the necessary modifications to include the additional new datasets.

We finish this section by noting that although in Table 1.1 we have given more than 10 parameters, almost all the currently existing observational data can be accounted for by a model having only 5 (6) freely adjustable parameters (Liddle 2004, Tegmark et al. 2004b):  $h$ ,  $\Omega_m$ ,  $\Omega_b$ ,  $A_s$ ,  $\tau$ , ( $n_s$ ) and the others just kept fixed to their “vanilla” values:  $\Omega_{DE} = 1 - \Omega_m$ ,  $\Omega_r = 1.68 \times \Omega_\gamma$ ,  $\Omega_\nu = 0$ ,  $\frac{d \ln n_s}{d \ln k} = 0$ ,  $A_t = 0$ ,  $w_{\text{eff}} = -1$ , ( $n_s = 1$ ). These simple 5 – 6 parameter models can be extended in several ways, but as already mentioned, current data does not really “call for any extra parameters”. Two interesting questions to be hopefully settled in the nearest future are:

- Is the dark energy equation of state  $w_{DE} = -1$  independent of redshift, i.e. is it compatible of being a cosmological constant?
- Do we live in a marginally closed Universe as might currently be hinted by the WMAP data?

---

<sup>23</sup> Assuming that we started out the chain in a high likelihood region.

<sup>24</sup> <http://cosmologist.info/cosmomc/>



## 1.4. Large galaxy/cluster surveys. Fast semianalytical methods for structure formation

Several galaxy/galaxy cluster redshift surveys have played an important role in helping to establish the “concordance” cosmological model. The largest redshift survey in existence is the Sloan Digital Sky Survey (SDSS)<sup>25</sup> with its latest public data release providing 565,715 galaxy and 76,483 quasar redshifts across 4783 square degrees of the sky. The Main Galaxy Sample reaches redshifts  $z \sim 0.2$ , whereas the subset of the spectroscopic galaxy sample known as the Luminous Red Galaxy (LRG) sample reaches redshifts  $z \sim 0.5$ . The SDSS is actually both imaging and redshift survey, where the imaging is taken in five photometric bands. Out of this high quality imaging data various classes of objects are selected (like Luminous Red Galaxies, quasars etc.) for subsequent spectroscopic follow-up. The survey, once finished, is planned to provide redshifts for  $\sim 10^6$  Main Sample galaxies,  $\sim 10^5$  LRGs,  $\sim 10^5$  quasars and cover  $\sim \frac{1}{4}$  of the sky. The imaging part of the survey has already covered the initially planned area. The spectroscopic part, for which  $\sim 80\%$  is currently completed, will be also finished within a couple of years in the framework of the recently announced SDSS II. Another large survey, which finished taking data in 2002, is the 2dF Galaxy Redshift Survey.<sup>26</sup> This survey covers  $\sim 2000$  square degrees and provides redshifts for  $\sim 220,000$  galaxies. The depth of the survey is comparable to the SDSS Main Sample, i.e.  $z \sim 0.2$ .

The redshift space distribution of galaxies can be used to estimate the power spectrum of the underlying matter distribution, which is sensitive to several cosmological parameters (see subsection 1.1.2). However, there are several complications involved:

1. Galaxies/clusters represent discrete sampling of the underlying matter density field. To be able to extract the clustering signal one needs sufficiently high sampling density in order to avoid the dominance of the shot noise component.
2. Galaxies/clusters do not faithfully trace the mass. In the CDM models the more massive the object, the more strongly clustered it is with respect to the underlying density field. This is known as biasing (e.g. [Kaiser 1984](#)). In general the bias can be scale-dependent, nonlinear and stochastic ([Dekel & Lahav 1999](#)). Only on the largest scales the usual approximation of linear scale-independent bias is justified.
3. The distances inferred from the measurements of the redshifts are distorted due to the non-negligible peculiar motions of the galaxies (the so-called redshift space distortions). Spatially extended coherent inflows of matter towards massive accretion centers boost the redshift space power spectrum over the real space one at large scales ([Kaiser 1987](#)). Inside the galaxy clusters the chaotic motion of galaxies causes the cluster to appear elongated along the line of sight, known as the “fingers of God” effect (e.g. [de Lapparent et al. 1986](#)). This results in the drop of power with respect to the real space power spectrum. Redshift space distortions somewhat complicate the power spectrum analysis, but luckily they also carry complementary cosmological information.

---

<sup>25</sup><http://www.sdss.org/>

<sup>26</sup><http://www.mso.anu.edu.au/2dFGRS/>

4. To convert the observed redshifts to comoving distances one has to assume some cosmological model. If the real cosmology differs from the chosen fiducial model, we end up with a distorted power spectrum. This is known as the cosmological distortion.
5. The theory makes accurate predictions only about the linearly evolved power spectrum, whereas with redshift surveys we are able to probe down to small scales that have gone nonlinear long ago. In order to exploit this extra information at smaller scales one needs to have some description available that goes beyond the simple linear evolution theory. Since currently our understanding of the nonlinear evolution is rather limited, one usually ignores completely the small-scale data above the wavenumbers of  $\sim 0.2 h \text{ Mpc}^{-1}$ .

The power spectrum of the SDSS main galaxy sample has been measured by Tegmark et al. (2004a). The cosmological parameter estimation based on this measurement along with the CMB data from the WMAP experiment was carried out in Tegmark et al. (2004b). The correlation function analysis for the SDSS LRG sample can be found in Zehavi et al. (2005), Eisenstein et al. (2005). The 2dF survey power spectrum has been determined by Percival et al. (2001) and more recently by Cole et al. (2005), the parameter estimation including additional CMB data was performed in Percival et al. (2002), and the spatial two-point correlation function was analyzed by Hawkins et al. (2003).

These surveys certainly contain a lot of other information beyond the simple power spectrum (correlation function) descriptor. Especially rich is the dataset provided by the SDSS with its high quality multicolor photometry along with a spectra that have far higher resolution than would just be needed to determine redshifts. The question is of course how much of this information we might possibly use for cosmological purposes, taking into account our currently rather limited understanding of the nonlinear structure formation. In this thesis we limit ourselves only to the two point descriptors of galaxy clustering. A detailed review of the various other statistical quantities which are sensitive to higher order correlations of the matter distribution can be found in Martínez & Saar (2002).

For the next generation of large redshift surveys to become operational one probably has to wait till  $\sim 2012$ . There is a WFMOS (Wide Field Multi-Object Spectrograph)<sup>27</sup> instrument construction planned for the Gemini<sup>28</sup> and Subaru<sup>29</sup> observatories that should be completed at 2012. This new multi-object spectrograph will be able to measure the spectra of  $\sim 5000$  objects at the same time. There are plans to perform a wide field ( $\sim 2000 \text{ deg}^2$ ) redshift survey giving spectra for  $\sim 2 \times 10^6$  galaxies up to redshifts of  $z \sim 1.3$  together with a more narrow ( $\sim 200 \text{ deg}^2$ ) and deeper ( $z \sim 2 \dots 3$ ) survey with a yield of  $\sim 5 \times 10^5$  galaxies. With the capabilities of this new instrument all of this could be obtained with an observational effort lasting significantly less than one year. By the year 2020 the field of large-scale structure studies is probably completely dominated by the proposed next-generation radio synthesis array, the Square Kilometre Array (SKA)<sup>30</sup>. With its  $1 \text{ km}^2$  collecting area, wide frequency coverage ( $0.1 - 25 \text{ GHz}$ ), and very large field of view (possibly up to  $100 \text{ deg}^2$ ) this revolutionary instrument would be able to measure within the framework of the observational campaign lasting  $\sim 1$  year the redshifts of

---

<sup>27</sup>Formerly known as K.A.O.S. (Kilo-Aperture Optical Spectrograph).

<sup>28</sup><http://www.gemini.edu/>

<sup>29</sup><http://www.naoj.org/>

<sup>30</sup><http://www.skatelescope.org/>

$\sim 10^9$  galaxies up to  $z \sim 1.5$  via their 21 cm line emission of neutral hydrogen.<sup>31</sup> With such a huge number of galaxy redshifts available one can find the 3D power spectrum and its evolution with extremely high accuracy. Being able to determine the locations of the acoustic features in the power spectrum for several redshifts, one can put very tight constraints on the properties of the dark energy. Although the next significant step in mapping the 3D distribution of galaxies will probably be made around  $\sim 2012$ , there is some short-term progress hoped by exploiting the new multi-object spectrographs such as AAOmega<sup>32</sup> (the successor of the 2dF instrument) on the Anglo-Australia Telescope<sup>33</sup> and FMOS<sup>34</sup> on the Subaru telescope. These instruments, if given sufficiently observational time, could provide hundreds of thousands of galaxy redshifts over large areas of sky within a few coming years. Also the SDSS survey will be completed around 2008, roughly doubling the size of the spectroscopic sample collected so far.

On the other hand, the next generation of wide-field imaging surveys are already operating or will become operational very soon. The wide-field optical cameras currently collecting data include OmegaCAM<sup>35</sup> on VST<sup>36</sup> and MegaCam<sup>37</sup> on Canada-France-Hawaii Telescope<sup>38</sup>. In the near infrared there is a WFCAM instrument<sup>39</sup> on United Kingdom Infra-Red Telescope<sup>40</sup> collecting data within the framework of the UKIDSS<sup>41</sup> infrared survey (the successor of 2MASS<sup>42</sup>). In late 2006 the construction of the 4 m class wide-field survey telescope VISTA (Visible and Infrared Survey Telescope for Astronomy)<sup>43</sup>, equipped with near infrared camera, will be finished in Chile. Within a few coming years the Pan-STARRS (Panoramic Survey Telescope & Rapid Response System)<sup>44</sup> survey with its four 1 m telescopes will start to map the sky. This facility is able to image the entire available sky several times each month (thus providing a breakthrough in detecting transient events). With this instrument a 1200 deg<sup>2</sup> ultra deep imaging survey is also planned. In the more distant future ( $\sim 2012$ ) the LSST (Large Synoptic Survey Telescope)<sup>45</sup> is planned to become operational. This instrument has capabilities to cover the entire available part of the sky in every three nights. It is estimated that from the weak lensing maps provided by LSST it should be possible to detect  $\sim 2 \times 10^5$  galaxy clusters. As the imaging of all these surveys is taken in several filters it is possible to use the photometric redshift (photo- $z$ ) techniques to obtain approximate estimates for the redshifts of the detected objects. To calibrate photo- $z$  one still needs complementary redshift surveys to measure the redshifts for some small fraction of the objects. With this effectively “2.5D data” one can certainly perform many useful studies of the large-scale structure. The lack of positional accuracy along the line of

---

<sup>31</sup>In 2D the number of detected galaxies would reach  $\sim 10^{10}$ .

<sup>32</sup><http://www.ast.cam.ac.uk/AAO/local/www/aaomega/>

<sup>33</sup><http://www.aao.gov.au/about/aat.html>

<sup>34</sup><http://www.sstd.rl.ac.uk/fmos/>

<sup>35</sup><http://www.astro.rug.nl/~omegacam/>

<sup>36</sup>[http://twg.na.astro.it/vst/vst\\_homepage\\_twg.html](http://twg.na.astro.it/vst/vst_homepage_twg.html)

<sup>37</sup><http://cfht.hawaii.edu/Instruments/Imaging/MegaPrime/>

<sup>38</sup><http://www.cfht.hawaii.edu/>

<sup>39</sup><http://www.roe.ac.uk/atc/projects/wfcam/>

<sup>40</sup><http://www.jach.hawaii.edu/UKIRT/>

<sup>41</sup><http://www.ukidss.org/>

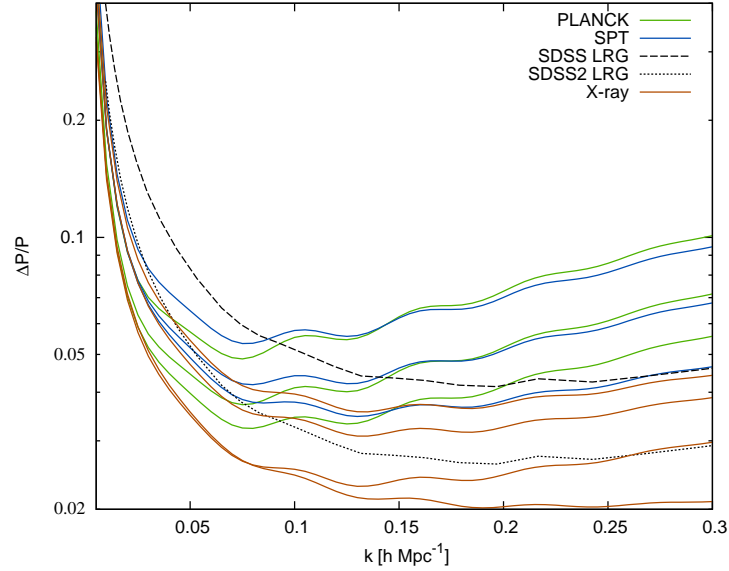
<sup>42</sup><http://www.ipac.caltech.edu/2mass/>

<sup>43</sup><http://www.vista.ac.uk/>

<sup>44</sup><http://pan-starrs.ifa.hawaii.edu/public/>

<sup>45</sup><http://www.lsst.org/>

sight can be substantially compensated by much larger sky coverage compared to the currently existing redshift surveys. Although being capable of providing significant new information in a relatively short timescale, these surveys by no means can replace the future wide-field deep redshift surveys.



**Figure 1.6.:** The performance of various cluster surveys in comparison to the SDSS LRG sample. For the SPT and PLANCK-like SZ surveys we have plotted the cases with 15,000, 25,000 and 35,000 detected galaxy clusters. The lines corresponding to the flux-limited X-ray survey represent the cases with 25,000, 35,000, 70,000 and 100,000 detected clusters. With the dotted lines we have also shown the obtainable accuracy of the power spectrum measurement once the SDSS redshift survey is completed within the few coming years. The PLANCK-like, SPT-like, and possible X-ray survey were assumed to cover the full sky, one octant, and 60% of the sky, respectively.

In the upcoming years also several cluster surveys will be performed. The most remarkable amongst these are the SPT (South Pole Telescope) <sup>46</sup> and PLANCK <sup>47</sup> surveys that exploit the specific frequency dependence and redshift independence of the thermal Sunyaev-Zeldovich effect (Sunyaev & Zeldovich 1980) for finding galaxy clusters. PLANCK as the next generation CMB satellite mission will provide nearly full sky coverage (except the regions close to the Galactic plane that are strongly “contaminated” by several foreground components), whereas the SPT is planned to cover  $\sim 4000 \text{ deg}^2$  of the sky. Due to the relatively poor angular resolution PLANCK will be able to detect only low redshift ( $z \sim 0.5$ ) galaxy clusters, while the SPT cluster sample will be much deeper ( $z \sim 1$  and beyond). Both of these surveys are expected to detect  $\sim 25,000$  galaxy clusters. In order to use the full power of these samples for large-scale structure studies, one needs to determine redshifts of these objects, which is a very difficult task in practice, since one needs at least  $\sim 10$  galaxy redshifts per cluster. On the other hand, it is

<sup>46</sup><http://spt.uchicago.edu/>

<sup>47</sup>[www.rssd.esa.int/Planck/](http://www.rssd.esa.int/Planck/)

also possible to use photo- $z$  instead. It is completely fine to use photo- $z$  for cluster number count studies, but for the clustering analysis this would lead to significant loss of (potential) information. As noted earlier, large cluster samples are similarly expected to be obtained by the weak lensing surveys. Also there are some plans in the X-ray community to perform a cluster survey with a yield of  $\sim 10^5$  galaxy clusters (Hasinger 2005). Although the cluster samples will be significantly sparser than the samples provided by the galaxy redshift surveys, the higher level of clustering strength of these massive systems will somewhat compensate the lower spatial density. In Fig. 1.6 we have shown the potential of the large cluster samples to constrain the power spectrum in comparison to the SDSS LRG sample. Here we have assumed that the spectrum will be determined in wavenumber bins of width  $\Delta k = 0.005 h \text{ Mpc}^{-1}$ . With the dotted line we have shown the performance of the SDSS LRG redshift survey once completed. The blue/green lines correspond to the PLANCK/SPT-like SZ cluster survey with 15,000, 25,000 and 30,000 detected galaxy clusters. The red lines represent the limits obtainable from the flux-limited X-ray survey with 25,000, 35,000, 70,000 and 100,000 galaxy clusters. The PLANCK-like, SPT-like, and possible X-ray survey were assumed to cover the full sky, one octant, and 60% of the sky, respectively. Here we remind that for the “concordance” cosmological model a typical predicted level of the fluctuations in the matter power spectrum  $\frac{\Delta P}{P} \sim 5\%$ .

In order to understand the influence of the cosmic variance, the survey selection effects and the redshift-space distortions on the statistical descriptors extracted from the realistic galaxy/cluster samples, one often has to rely on the Monte Carlo approach that frequently requires us to build many thousands of mock catalogs over large cosmological volumes, before converging results are obtained. Moreover, changing the underlying assumed cosmological model, one should, in principle, redo the whole calculation. With current computational technology it is completely infeasible to attack these issues using numerically demanding N-body calculations. So inevitably one has to relay on fast semianalytical techniques. Many years of work on N-body calculations has given us an accurate description for the mass function of the CDM haloes (Sheth & Tormen 1999, Jenkins et al. 2001, Sheth et al. 2001). It would also be desirable to have some fast tools available that are capable to reliably generate the spatial distribution of haloes along the past light-cone. Here we just mention a few methods that might be useful to achieve these goals.

The evolution of the CDM component, which governs the formation of the large-scale structure<sup>48</sup>, is determined by the Vlasov<sup>49</sup>-Poisson equation system. Since this is a complicated system to deal with, one usually makes a fluid approximation, which is valid at early enough times. Modeling the CDM as a pressureless fluid leads to the replacement of the Vlasov-Poisson system with a simpler system that consists of the continuity equation, the Euler equation and the Poisson equation. At early enough time, when the density fluctuations  $\delta \ll 1$ , one can linearize these equations. This leads to the standard Eulerian linear perturbation theory (see e.g. Peebles 1980, Sahni & Coles 1995). One can also go to higher perturbative orders (e.g. Peebles 1980, Fry 1984, Sahni & Coles 1995, Bernardeau et al. 2002). If the fluctuation amplitude  $\delta \sim 1$ , the perturbative approach breaks down completely. It turns out that much better approximations that stay valid also in the quasi-linear regime can be obtained by using the particle-based Lagrangian perturbative approach (see e.g. Moutarde et al. 1991, Bouchet et al. 1992, Buchert 1992, Bouchet

<sup>48</sup>The influence of baryons i.e. gasodynamical effects become noticeable only on smaller scales relevant e.g. for the galaxy formation.

<sup>49</sup>Also known as collisionless Boltzmann equation.

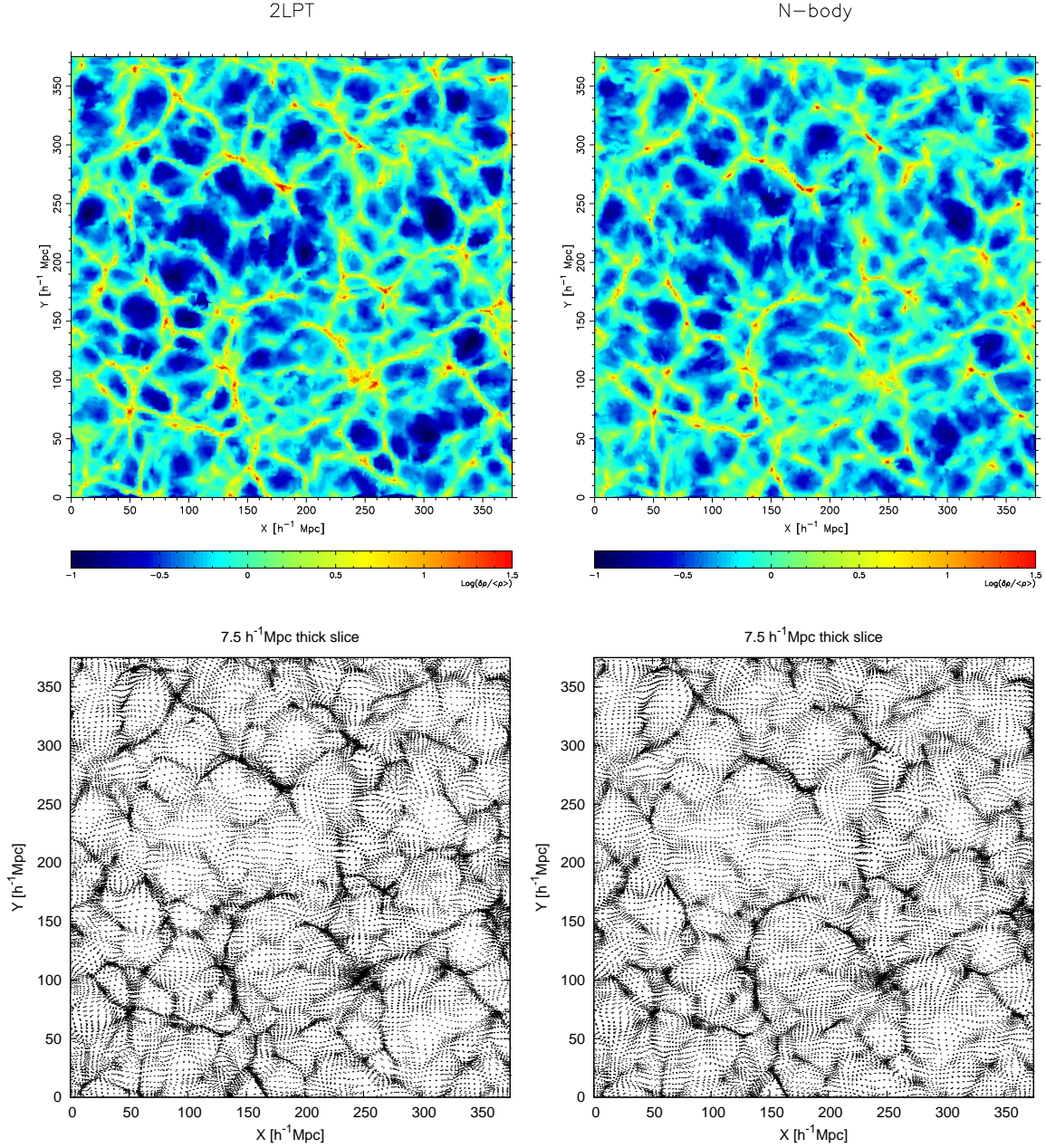


et al. 1995, Sahni & Coles 1995). Here instead of perturbing the density contrast field  $\delta$ , one perturbs particle trajectories (i.e. gives a description in terms of the displacement field). The lowest order Lagrangian approach is known as the Zeldovich approximation (Zel'Dovich 1970)<sup>50</sup> In this approximation the particles just keep on moving with their initial comoving velocities. It provides a good description of the evolution of the density field up to the formation of the first caustic structures. After that the approximation gets progressively worse since the ballistically moving particles do not feel any gravitational pull of the dense caustic regions. Instead they just pass through these density enhancements, leading to their complete erasure at later times. This method can be modified by including extra viscosity term to the equations, which makes the particles to stick together once they enter into caustics— and thus the name— the adhesion approximation (Gurbatov et al. 1989). There is actually even easier method available to cure the “caustic dilution problem”. It turns out that if one removes from the initial displacement field the modes shorter than the nonlinearity scale corresponding to the redshift when the “output is taken”, the caustics remain sharp. This is known as the truncated Zeldovich approximation (Coles et al. 1993, Melott et al. 1994). The optimal choice for the spectral truncation scale has been investigated in several papers e.g. Coles et al. (1993), Melott et al. (1994), Weiss et al. (1996), Hamana (1998). As with an Eulerian approach here one can also work out the higher order perturbation theory. In this thesis we use the second order “optimized” (i.e. optimally truncated) Lagrangian perturbation theory (2LPT) to generate large-scale density fields, which turn out to agree (at scales larger than  $\sim 10 h^{-1}$  Mpc) remarkably well with those obtained from direct N-body simulations started from the same initial conditions. For the description of the exact realization of the method see Appendix C. In Fig. 1.7 we compare the performance of the 2LPT with respect to the full N-body calculation. Both calculations had  $256^3$  dark matter particles in a comoving box of side length  $375 h^{-1}$  Mpc. The N-body calculation used the MLAPM<sup>51</sup> (Multi-Level Adaptive Particle Mesh) cosmological multi-grid N-body solver (Knebe et al. 2001). In the upper panel of Fig. 1.7 we show the cuts through the 3D density fields, constructed with the 3D Delaunay tessellation field estimator as described in Schaap & van de Weygaert (2000), whereas the lower panels display the corresponding particle distributions inside the  $7.5 h^{-1}$  Mpc thick slices at the position of the cuts. As we can see, at the current resolution level the 2LPT density field is remarkably similar to the one obtained from the full N-body calculation. Now the biggest issue is of course how to fragment this density field into separate haloes. With the cluster-size haloes this task is probably relatively easily manageable, since these massive systems can be directly identified with the density peaks where several filamentary features cross. The masses of these peaks (e.g. determined with a  $10 h^{-1}$  Mpc top-hat filter) can be rank-ordered and then mapped to the “real cluster masses” compatible with the analytically well known cluster mass function. With the galaxies the situation is much more complicated since the formation of these smaller scale systems involves a great deal of gasodynamical processes, which are poorly understood. Of course, the approximation based on the 2LPT breaks down already on scales way larger than the ones relevant for the galaxy formation. Here one can take a completely statistical approach instead, and assume that the probability to form a galaxy at some point is determined by the density of its immediate neighborhood. In the simplest case this probability can just be taken to

---

<sup>50</sup>In fact the Zeldovich approximation provides an exact solution for the evolution of the plane-wave perturbation in the expanding background up to the first shell-crossing.

<sup>51</sup><http://www.aip.de/People/AKnebe/MLAPM/>



**Figure 1.7.:** The performance of the “optimized” 2LPT (left-hand panels) with respect to the precise N-body (right-hand panels) calculation. The computational box had a side length of  $375 h^{-1}$  Mpc and contained  $256^3$  dark matter particles. The upper panels show the cuts through the 3D density fields, constructed with the 3D Delaunay tessellation field estimator as described in [Schaap & van de Weygaert \(2000\)](#), whereas the lower panels display the corresponding particle distributions inside the  $7.5 h^{-1}$  Mpc thick slices at the position of the cuts.

be proportional to the surrounding density. As the initial density field itself was a realization of the Gaussian random process, and the probability for forming a galaxy at some point is modeled as a Poisson process with the intensity given by the density of the underlying field, we have here a doubly random process, which in spatial statistics is known as a Cox process (see e.g. [Martínez & Saar 2002](#)). Actually almost all the statistical analyses of the large-scale structure assume that the spatial point distribution provided by the galaxies can be modeled as a Cox process i.e. the galaxies represent a Poisson sampling of the underlying density field. One can make this process slightly more realistic, for example by adding around each “formed galaxy” a zone of avoidance, where other galaxies cannot form. In this thesis we use this approach to generate mock galaxy catalogs needed for determining the covariance properties of the measured SDSS LRG power spectrum.

## 1.5. In this Thesis

In this introductory Chapter we have summarized the underlying ideas and have hopefully helped to set the stage for what follows.

In Chapter 2 we investigate the possibility of using the forthcoming blank sky wide-field SZ-selected cluster surveys to put constraints on the properties of dark energy. Instead of the usual number counts analysis we concentrate on the clustering properties of those largest gravitationally relaxed systems. We carry out Fisher matrix forecasting for various cosmological parameters, specifically focusing on two types of survey: (i) shallow surveys with approximately full sky coverage, like the one to be performed by the PLANCK Surveyor spacecraft, (ii) narrower ( $\sim 1/8$  of the sky) and deeper surveys, e.g. the one that will be carried out by the SPT. The analysis includes a treatment of the anisotropic nature of the observed power spectra due to redshift space and cosmological distortions. The light-cone effect is also taken into account. We pay particular attention to the possibility of detecting baryonic acoustic oscillations in the cluster power spectrum. It turns out that on the largest scales both of these surveys will have the capability to improve measurement of acoustic features compared to the recently announced results using the SDSS LRG ([Eisenstein et al. 2005](#), [Hütsi 2005](#), [Hütsi 2006a](#)) and 2dF ([Cole et al. 2005](#)) galaxy samples. We demonstrate that interesting constraints for the dark energy equation of state parameter can be expected if additional information from the CMB studies is included.

In Chapter 3 we carry out a full redshift-space power spectrum analysis of the SDSS DR4 LRG sample, finding evidence of acoustic oscillations down to scales of  $\sim 0.2 h \text{ Mpc}^{-1}$ , which approximately corresponds to the seventh peak in the CMB angular power spectrum. We present a high-precision measurement of the acoustic scale:  $(105.4 \pm 2.3) h^{-1} \text{ Mpc}$ . Using the WMAP data together with the prior on the Hubble parameter from the HST Key Project,  $H_0 = 72 \pm 8 \text{ km/s/Mpc}$ , the corresponding scale would be predicted to lie in the range  $(107 \pm 20) h^{-1} \text{ Mpc}$ , showing that our measurement provides an order of magnitude improvement over that prediction. We also show that this measurement is in good agreement with the correlation function analysis of the SDSS LRGs carried out by [Eisenstein et al. \(2005\)](#). It is additionally demonstrated that models with baryonic features in their spectra are favored by  $3.3\sigma$  over their “smoothed-out” counterparts without any oscillatory behavior, i.e. the acoustic features are detected at a relatively high confidence level.

In Chapter 4 we work out the consequences of the measured power spectrum and acoustic scale



on various cosmological parameters. In this parameter study we focus on adiabatic, spatially flat models with negligible massive neutrino and tensor perturbation contributions. To break the parameter degeneracies additional data from the WMAP experiment is used. The most remarkable result is a new determination of the Hubble parameter  $H_0 = 70.8^{+1.9}_{-1.8}$  km/s/Mpc. This precise measurement of  $H_0$  helps to break several parameter degeneracies and allows us to measure the density parameters  $\Omega_{cdm}$  and  $\Omega_b$ , and also the dark energy equation of state parameter  $w_{\text{eff}}$  with significantly higher accuracy than is possible with the WMAP and HST data alone. Through the determination of these parameters we are able to constrain the low-redshift expansion law of the Universe. Particularly, we find that the decelerating Universe is ruled out at the confidence level of  $5.5\sigma$ . We also stress the need to properly rescale the power spectrum if the background model is changed. Surprisingly, almost all previous analyses have ignored this point.



## 2. Clustering of SZ clusters on a past light-cone: acoustic oscillations and constraints on dark energy

G. Hütsi

*Astronomy & Astrophysics 446, 43 (2006)*

### Abstract

We study the clustering of SZ-selected galaxy clusters on a past light-cone, particularly paying attention to the possibility of constraining properties of dark energy. The prospects of detecting baryonic features in the cluster power spectrum for a wide and shallow survey like PLANCK, and for an SPT-like narrow and deep survey are discussed. It is demonstrated that these future blank sky SZ surveys will have the capability to improve over the recently announced detection of baryonic oscillations based on the SDSS Luminous Red Galaxy (LRG) sample. We carry out parameter estimation using a Fisher matrix approach taking into account the anisotropic nature of the power spectrum due to redshift space and cosmological distortions. The clustering signal which is not too sensitive to systematic uncertainties serves as a valuable piece of information that in combination with other sources of data helps in breaking degeneracies between the cosmological parameters.

### 2.1. Introduction

In the early 1970's it was recognized that acoustic waves in the radiation dominated matter prior to the epoch of recombination of hydrogen in the Universe resulted in the characteristic pattern of maxima and minima in the post-recombination matter power spectrum (Sunyaev & Zeldovich 1970, Peebles & Yu 1970, Doroshkevich et al. 1978). These acoustic peaks depend on the size of the sound horizon and on the relative phases of the perturbations containing different masses at the moment of recombination. In the currently most favorable cosmological models with a Cold Dark Matter (CDM) component dominating significantly over the baryonic part, the acoustic features in the matter power spectrum are strongly damped, reaching only  $\sim 5\%$  level for the “concordance” model (Bahcall et al. 1999, Spergel et al. 2003). For early description of the acoustic oscillations in the context of the CDM models see Blumenthal et al. (1988). For a given set of parameters of the Universe ( $\Omega_b h^2$ ,  $\Omega_m h^2$ ) the position of the maxima and minima are fully determined (e.g. Eisenstein & Hu 1998). Acoustic oscillations also leave their imprint on the angular perturbations of the Cosmic Microwave Background (CMB). MAXIMA-

1, <sup>1</sup> Boomerang, <sup>2</sup> WMAP, <sup>3</sup> VSA, <sup>4</sup> CBI, <sup>5</sup> and many other CMB experiments detected the first acoustic peaks in the CMB power spectrum with a high confidence level (Hanany et al. 2000, Netterfield et al. 2002, Bennett et al. 2003, Grainge et al. 2003, Pearson et al. 2003). These observations gave very important information about the key parameters of the Universe using the angular scale of acoustic features as rulers and taking into account the ratio of amplitudes of the different peaks.

The existence of acoustic rulers in the Universe is of enormous importance since they permit us to measure the behavior of the Hubble parameter with redshift and also allow us to establish the distance-redshift relation. This is especially important now when the discovery of *dark energy* (DE) is introducing more questions than the answers it provides. Different areas of observational cosmology (SNe Ia e.g. SCP, <sup>6</sup> high- $z$  SN search; <sup>7</sup> large scale structure surveys e.g. SDSS, <sup>8</sup> 2dF <sup>9</sup>; CMB experiments) provide evidence that the expansion of our Universe has been proceeding in an accelerated fashion since  $z \sim 0.75$  (Riess et al. 1998, Perlmutter et al. 1999, Spergel et al. 2003). Currently there is no physical understanding or even a reliable model for the DE. One of the first tasks to understand the nature of DE will be the measurement of its equation of state parameter,  $w = P/\rho$ , and its possible evolution with time  $w(z)$ . When we are equipped with a good standard ruler and are able to measure its angular behavior with redshift, we can obtain very valuable information about  $w(z)$ . With the CMB data we can determine an angular diameter distance to the last scattering surface with high precision. By combining this information with the measurement of acoustic peaks in the distribution of baryons at lower redshifts  $0 < z < 1 - 2$ , we will have unique information about the effective  $w$  and may even be able to determine  $w(z)$  (Eisenstein & Hu 1998, Blake & Glazebrook 2003, Hu & Haiman 2003, Linder 2003, Seo & Eisenstein 2003).

It was obvious since the first publications that acoustic oscillations should also leave their imprint on the large scale structure of the Universe and thus influence e.g. the correlation function and the power spectrum of galaxies. The first successful detection of these features was presented by Eisenstein et al. (2005) who found traces of acoustic oscillations in the distribution of luminous red galaxies for which they had excellent measurements of angular positions and redshifts obtained by the SDSS collaboration.

In this paper we discuss the opportunities that will be opened by the planned blank sky deep SZ cluster surveys which will be performed in the coming years. Clusters of galaxies are especially interesting objects for the study of acoustic features in the spatial distribution of objects since it has long been known that *the clustering of clusters is an order of magnitude enhanced in comparison to galaxies* (Bahcall & Soneira 1983, Kaiser 1984). Therefore even with smaller statistics of clusters it is possible to get useful results.

In the next few years there will be very deep SZ cluster surveys of the restricted regions of

---

<sup>1</sup><http://cosmology.berkeley.edu/group/cmb/>

<sup>2</sup><http://cmb.phys.cwru.edu/boomerang/>

<sup>3</sup><http://map.gsfc.nasa.gov>

<sup>4</sup><http://www.mrao.cam.ac.uk/telescopes/vsa/>

<sup>5</sup><http://www.astro.caltech.edu/~tjp/CBI/>

<sup>6</sup><http://panisse.lbl.gov>

<sup>7</sup><http://cfa-www.harvard.edu/cfa/oir/Research/supernova/HighZ.html>

<sup>8</sup><http://www.sdss.org/>

<sup>9</sup><http://www.aao.gov.au/2df>

the sky performed by several projects, e.g. APEX,<sup>10</sup> SZA,<sup>11</sup> AMI,<sup>12</sup> ACT<sup>13</sup>. Our analysis showed that the volume of these surveys and the number of possible cluster detections will be unfortunately insufficient for the search for the acoustic wiggles in the power spectrum. However, two planned surveys which will be carried out by the PLANCK Surveyor<sup>14</sup> spacecraft and the South Pole Telescope<sup>15</sup> (SPT) have very good prospects for the detection of acoustic features. PLANCK will make a shallow blank sky cluster survey permitting one to detect up to 20,000 rich clusters of galaxies (e.g. Majumdar & Mohr 2004) with the bulk of objects at  $z < 0.5$ , but will reach distances of  $z \sim 0.8$ . In contrast the SPT survey will observe deeper, but will cover only 10% of the sky. This survey is expected to detect up to 30,000 clusters of galaxies (Majumdar & Mohr 2004), and many of them will be at significantly higher redshifts compared to the ones observed by PLANCK. Unfortunately it is not enough to measure only the SZ flux or brightness of the clusters. In order to measure the equation of state of DE *we need the redshift estimate for each cluster in the sample* which will be hard and time consuming work for many optical, X-ray and possibly radio astronomers. However, when this problem is solved, cosmologists will have a unique sample of clusters of galaxies with good knowledge of their angular position, redshift, and hopefully also mass. In this paper we investigate what limits to the DE equation of state might be obtained when these large experimental efforts are completed. It is obvious that in parallel other ways to measure  $w(z)$  will be implemented, but any additional and independent information will be useful. Especially important is that PLANCK and SPT surveys of clusters of galaxies will be performed in any case. Certainly, for many various purposes: (i) study of the redshift distribution of clusters, (ii) study of the properties of the clusters as a population, (iii) search for high- $z$  clusters etc., we always need to estimate redshift. Therefore the information on the power spectrum of clusters of galaxies, acoustic wiggles and the subsequent determination of  $w$  is complementary but extremely important part of these surveys.

The structure of this paper is as follows. In Sec. 2 we describe an analytical model for a cluster power spectrum on a light-cone and calibrate it against the numerical simulations. Sec. 3 discusses the possibility of detecting baryonic oscillations with the forthcoming SZ surveys. In Sec. 4 we carry out parameter forecasting using a Fisher matrix approach and Sec. 5 contains our conclusions.

## 2.2. Light-cone power spectrum of galaxy clusters

In this section we present the theoretical model for calculating the cluster power spectrum on our past light-cone. In order to assess the accuracy of the theoretical description, we make comparisons with the VIRGO Hubble Volume simulation outputs. We start with a very brief description of the VIRGO simulations and proceed with the calculation of power spectra using light-cone cluster catalogs provided by the VIRGO Consortium<sup>16</sup>.

---

<sup>10</sup><http://bolo.berkeley.edu/apexsz/>

<sup>11</sup><http://astro.uchicago.edu/sza/>

<sup>12</sup><http://www.mrao.cam.ac.uk/telescopes/ami/>

<sup>13</sup><http://www.hep.upenn.edu/act>

<sup>14</sup><http://astro.estec.esa.nl/Planck>

<sup>15</sup><http://astro.uchicago.edu/spt>

<sup>16</sup><http://www.mpa-garching.mpg.de/Virgo/>

### 2.2.1. Cluster power spectra from VIRGO simulations

We use outputs from the  $\Lambda$ CDM Hubble Volume simulation that was run in a  $3000 h^{-1}$  Mpc comoving box with a particle mass of  $2.25 \cdot 10^{12} h^{-1} M_{\odot}$ . The other simulation parameters were as follows:  $\Omega_m = 0.3$ ,  $\Omega_{\Lambda} = 0.7$ ,  $\Omega_b = 0.04$ ,  $h = 0.7$  and  $\sigma_8 = 0.9$  (for further details see e.g. [Evrard et al. 2002](#)). In our study we used  $z = 0$  snapshot, SphereB and OctantB light-cone catalogs<sup>17</sup>. The  $z = 0$  cluster catalog used a friend-of-friend scheme with a linking length  $b = 0.164$  for cluster identification, while for the light-cone outputs the spherical overdensity method with the overdensity 200 relative to the critical density was applied. The minimum number of particles per cluster is 30 and 12 for the  $z = 0$  and light-cone catalogs, respectively.

To calculate the power spectrum we follow the direct method of [Feldman et al. \(1994\)](#) (FKP), which is shown to be optimal for sufficiently large  $k$ -modes i.e.  $k \gg 1/L$ , where  $L$  is the typical spatial extent of a survey volume ([Tegmark et al. 1998](#)). Because FFTs are used to achieve significant speedup for Fourier sum calculations, we first have to find the density field on a grid. To this end we use the Triangular Shaped Cloud (TSC) ([Hockney & Eastwood 1988](#)) mass assignment scheme. Thus our density field is a filtered version of the underlying field, and as shown in [Jing \(2005\)](#), the real power spectrum  $P$  can be expressed as the following sum over aliases (correct again for the case  $k \gg 1/L$ ):

$$P_{\text{raw}}(\mathbf{k}) \simeq \sum_{\mathbf{k}'} S^2(\mathbf{k}') \sum_{\mathbf{n} \in \mathbb{Z}^3} \mathcal{W}^2(\mathbf{k} + 2k_N \mathbf{n}) P(\mathbf{k} + 2k_N \mathbf{n}) + \frac{1}{N} \sum_{\mathbf{n} \in \mathbb{Z}^3} \mathcal{W}^2(\mathbf{k} + 2k_N \mathbf{n}), \quad (2.1)$$

where the raw power spectrum:

$$P_{\text{raw}}(\mathbf{k}) \equiv \langle |\delta_g(\mathbf{k})|^2 \rangle \quad (2.2)$$

and Fourier transform of the overdensity field on a grid is calculated as usual:

$$\delta_g(\mathbf{k}) = \frac{1}{N} \sum_g \left[ n_g(\mathbf{r}_g) - \bar{n} S(\mathbf{r}_g) \right] e^{i \mathbf{r}_g \cdot \mathbf{k}}. \quad (2.3)$$

Here  $n_g$  is the number density field on a grid without any selection effect corrections,  $S$  is a selection function that also incorporates survey geometry (i.e.  $S = 0$  outside of survey boundaries),  $\bar{n}$  is the mean underlying number density and the sum runs over all grid cells. A Fourier transform of the selection function  $S(\mathbf{k})$  in Eq. (2.1) is normalized so that  $S(\mathbf{0}) = 1$  and the mass assignment window in the case of the TSC scheme can be expressed as:

$$\mathcal{W}(\mathbf{k}) = \left[ \frac{\prod_i \sin(\frac{\pi k_i}{2k_N})}{\prod_i \frac{\pi k_i}{2k_N}} \right]^3. \quad (2.4)$$

The second term on the left-hand side of Eq.(2.1) is the shot noise contribution and in the case of the TSC filter can be shown to give the following result ([Jing 2005](#)):

$$\text{SN} = \frac{1}{N} \prod_i \left[ 1 - \sin^2\left(\frac{\pi k_i}{2k_N}\right) + \frac{2}{15} \sin^4\left(\frac{\pi k_i}{2k_N}\right) \right]. \quad (2.5)$$

Summarized very briefly, our power spectrum calculation consists of the following steps:

<sup>17</sup>For the exact description of these catalogs see [Evrard et al. \(2002\)](#)

1. Determination of the selection function  $S$  (including survey geometry) and mean underlying number density  $\bar{n}$ ,
2. Calculation of the overdensity field on a grid using the TSC mass assignment scheme and its Fourier transform as given in Eq. (2.3),
3. Subtraction of the shot noise term (Eq. (2.5)) from the raw power spectrum (Eq. (2.2)),
4. Isotropization of the shot noise corrected power spectrum, i.e. averaging over  $k$ -space shells,
5. Application of normalization correction due to selection effects, i.e. dividing by  $\sum_{\mathbf{k}} S^2(\mathbf{k})$ ,
6. Deconvolving the smearing effect of the TSC mass assignment.

The “sharpening” in the last step is done using an iterative method as described in [Jing \(2005\)](#) with the only difference that here we do not approximate the power spectrum simply with a power law, but also allow for a running of the spectral index, i.e. we approximate it with a parabola in log-log coordinates.

When calculating the power spectrum in the above described way we assumed that the influence of selection/survey geometry effects on the power spectrum are separable. This is not the case for large scales ( $k \sim 1/L$ ) and also if too narrow (i.e.  $\Delta k \lesssim 1/L$ ) power spectrum bins are used. In the following we always make the power spectrum binning broad enough so that the neighboring bins can be safely assumed to be uncorrelated.

The power spectrum error is estimated using the simple “mode counting” result of FKP (see also [Tegmark et al. 1998](#)):

$$\frac{\Delta P}{P} = \sqrt{\frac{2}{V_{\text{eff}} V_k}}, \quad (2.6)$$

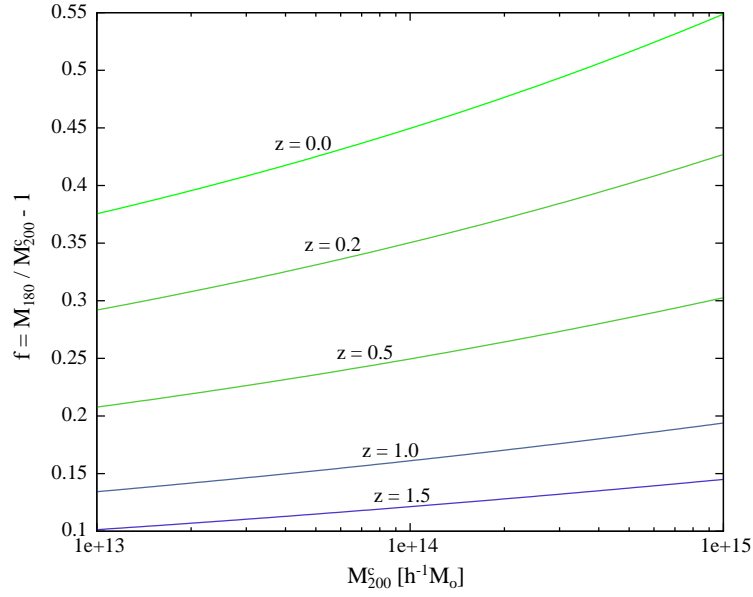
where  $V_k = 4\pi k^2 \Delta k / (2\pi)^3$  is the volume of the  $k$ -space shell and  $V_{\text{eff}}$  is the effective volume given by:

$$V_{\text{eff}} = \frac{\left[ \int W^2(z) \frac{dV_c}{dz} dz \right]^2}{\int W^4(z) \left[ 1 + \frac{1}{\bar{n}(z)P} \right]^2 \frac{dV_c}{dz} dz}. \quad (2.7)$$

Here  $dV_c$  is a comoving volume element and the weight function:

$$W(z) \propto \begin{cases} \text{const} & \text{for volume weighting} \\ \bar{n}(z) & \text{for number weighting} \\ \frac{\bar{n}(z)}{1 + \bar{n}(z)P} & \text{for an optimal FKP weighting.} \end{cases} \quad (2.8)$$

The halos of the VIRGO Hubble Volume simulations were identified using the spherical overdensity algorithm with an overdensity of 200 with respect to the critical density at the identification epoch ([Evrard et al. 2002](#)) (we denote the corresponding mass  $M_{200}^C$ ), while detailed comparisons suggest that the Press-Schechter type of analytical calculations ([Press & Schechter 1974](#), [Mo & White 1996](#), [Sheth & Tormen 1999](#), [Sheth et al. 2001](#)) provide a good match to simulations if an overdensity 180 with respect to the background density is used (corresponding



**Figure 2.1.:** Mass conversion from  $M_{200}^C$  to  $M_{180}$  for various redshifts.

mass  $M_{180}$ ) (e.g. [Jenkins et al. 2001](#)). In order to convert from one mass definition to the other we assume that the density profile of clusters is given by the NFW ([Navarro et al. 1997](#)) profile:

$$\rho(r) = \frac{\rho_s}{\left(\frac{r}{r_s}\right) \left(1 + \frac{r}{r_s}\right)^2}, \quad (2.9)$$

and the concentration parameter  $c = r_v/r_s$  ( $r_v$ -virial radius) and its evolution as a function of virial mass  $M_v$  is given as follows ([Bullock et al. 2001](#)):

$$c(M_v) = \frac{9}{1+z} \left( \frac{M_v}{M_*(z=0)} \right)^{-0.13}. \quad (2.10)$$

Here  $M_*$  is a standard nonlinear mass scale defined through  $\sigma(M_*, z) \equiv \delta_c(z)$ , where  $\sigma^2(M, z)$  is the variance of the linearly evolved density field on the comoving scale corresponding to the mass  $M$  at redshift  $z$  and  $\delta_c(z)$  is the spherical collapse threshold e.g. in Einstein-de Sitter model  $\delta_c(0) = 1.686$ .

Then the mass within the radius  $r$  can be expressed as:

$$M(< r) = 4\pi\rho_s r^3 f\left(\frac{r_s}{r}\right) = \frac{4\pi r^3}{3} \Delta \Omega_m \rho_c, \quad (2.11)$$

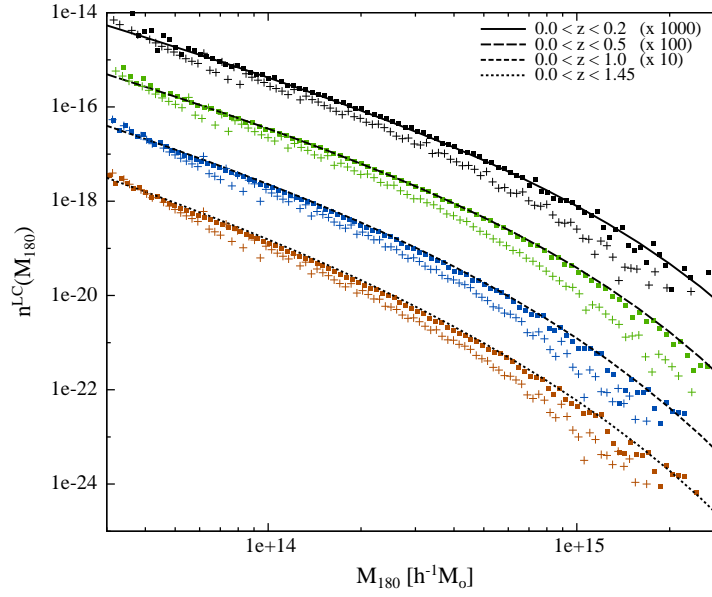
where

$$f(x) = x^3 \left[ \ln\left(1 + \frac{1}{x}\right) - \frac{1}{1+x} \right], \quad (2.12)$$

$\rho_c$  is the critical density and  $\Delta$  is the halo overdensity with respect to the background matter density at the epoch of halo identification.

In order to convert halo mass  $M$  corresponding to the overdensity  $\Delta$  to the one corresponding to the overdensity  $\Delta'$  we proceed as follows:





**Figure 2.2.:** Light-cone mass functions for different redshift intervals. For clarity the curves have been shifted by the factors given in the legend. Boxes show simulation results with the applied mass conversion from  $M_{200}^C$  to  $M_{180}$  whereas crosses are the results without any conversion.

1. From Eq. (2.11) determine the radius  $r$  corresponding to the mass  $M$  and overdensity  $\Delta$ ,
2. Solve

$$\Delta_v f\left(\frac{r_v}{c(r_v)r}\right) - \Delta f\left(\frac{1}{c(r_v)}\right) = 0 \quad (2.13)$$

for virial radius  $r_v$ . Here  $\Delta_v$  is the virial overdensity that we find numerically solving the spherical tophat collapse model (fitting formulae for  $\Delta_v$  for some cosmological models are given in [Bryan & Norman \(1998\)](#)),

3. Solve

$$\Delta_v f\left(\frac{r_v}{c(r_v)r'}\right) - \Delta' f\left(\frac{1}{c(r_v)}\right) = 0 \quad (2.14)$$

for  $r'$ ,

4. From Eq. (2.11) find  $M'$  corresponding to the radius  $r'$  and overdensity  $\Delta'$ .

The results of this mass conversion from  $M_{200}^C$  to  $M_{180}$  are shown in Fig. 2.1 as a fractional increase in mass  $f = M_{180}/M_{200}^C - 1$  for different redshifts. We see that especially for low redshift clusters this mass change can reach up to 50%. In Fig. 2.2 we demonstrate the importance of the mass conversion in order to get agreement with the analytical mass function calculations. Here

the mass function on a light-cone was calculated as follows:

$$n^{\text{LC}}(M_{180}, < z) = \frac{\int_0^z n(M_{180}, z) \frac{dV_c}{dz} dz}{\int_0^z \frac{dV_c}{dz} dz}. \quad (2.15)$$

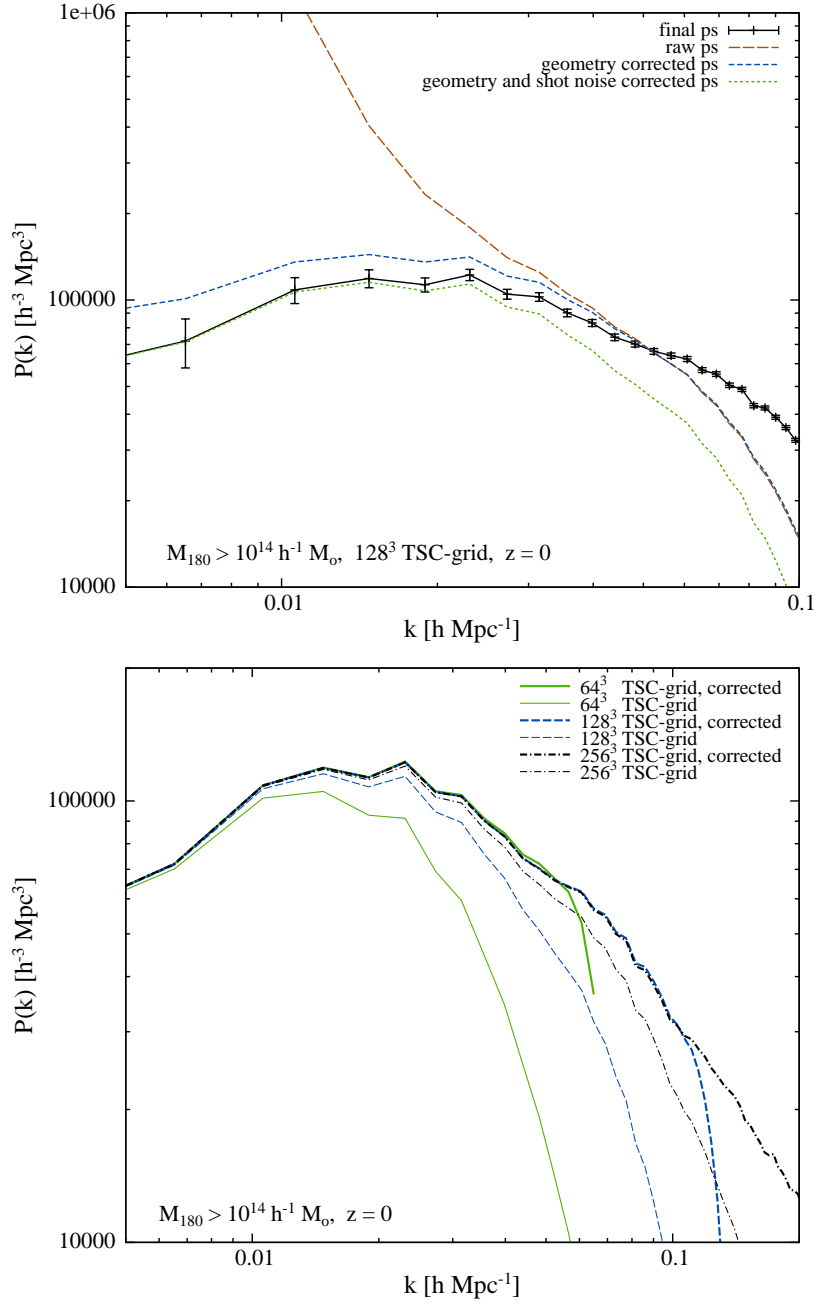
Here  $dV_c$  is a comoving volume element and  $n(M, z)$  is a mass function as described in [Sheth & Tormen \(1999\)](#), which is known to give a very good description of N-body results ([Jenkins et al. 2001](#)).

In the upper panel of Fig. 2.3 we show various corrections needed to achieve a reliable estimate of the power spectrum of the underlying cluster distribution while the lower panel demonstrates the consistency of the applied “sharpening” scheme. Here we have used the  $z = 0$  cluster catalog to allow for a comparison with the results presented in [Colberg et al. \(2000\)](#). The lower mass for the cluster selection was taken to be  $1.0 \cdot 10^{14} h^{-1} M_\odot$  in order to get the total number of objects equal to  $\sim 915,000$  as was used in [Colberg et al. \(2000\)](#). Also we have selected a spherical volume out of the full box to test how well the geometry correction works. The results of this comparison are given in Fig. 2.4. We see that the [Colberg et al. \(2000\)](#) power spectrum agrees with our calculations at the largest scales; however, for the smaller scales it drops below our results. We suspect that their correction for the grid smoothing effect was insufficient, although in their paper they do not describe how the power spectrum was calculated. As can be seen from the figure the shape of our cluster power spectrum agrees very well with the linear theory matter power spectrum up to the scale  $k \sim 0.15 h \text{ Mpc}^{-1}$ . Clearly with such a large number of clusters ( $\sim 477,000$  inside our spherical volume) baryonic oscillations are easily detectable and the corresponding “smooth” model without them is disfavored. The theoretical matter power spectra were calculated as described in [Eisenstein & Hu \(1998\)](#). Using the  $z = 0$  cluster catalog we also calculate power spectra and two-point correlation functions for various lower mass cutoffs. These results are presented in Fig. 2.5 where the left-hand panels show power spectra divided by the smooth model without baryonic oscillations and right-hand panels the respective correlation functions. Here the uppermost power spectrum is the same as the one given in Fig. 2.4. Solid/dotted lines show theoretical models with/without baryonic oscillations. Correlation functions were calculated using the estimator given by [Landy & Szalay \(1993\)](#):

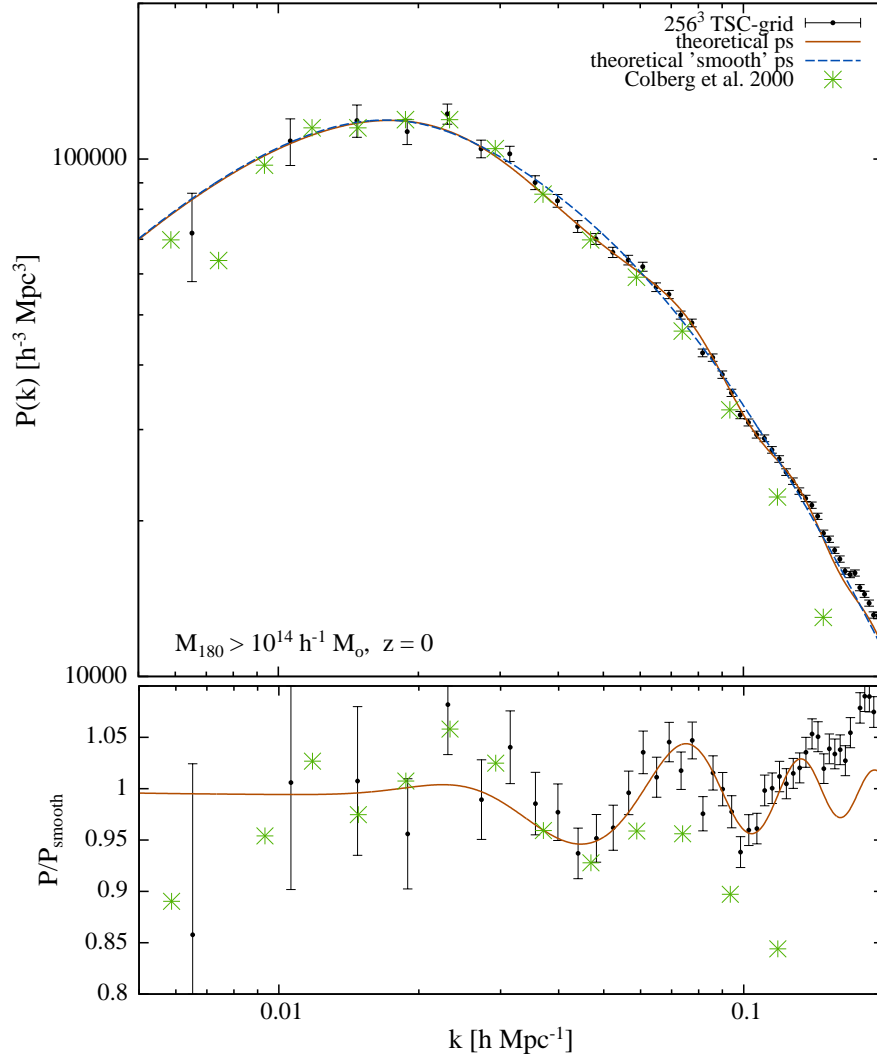
$$\xi(r) = \frac{DD - 2DR + RR}{RR}, \quad (2.16)$$

which has minimal variance for a Poisson process. Here DD, DR and RR represent the respective normalized data-data, data-random and random-random pair counts in a given distance range. Random catalogs were generated with ten times the number of objects in the main catalogs. The survey geometry was again taken to be a spherical volume reaching redshift  $z = 0.58$ . The number of objects corresponding to the lower mass cutoffs of  $1.0 \cdot 10^{14}$ ,  $2.0 \cdot 10^{14}$ ,  $3.0 \cdot 10^{14}$ ,  $4.0 \cdot 10^{14}$  and  $5.0 \cdot 10^{14} h^{-1} M_\odot$  (friend-of-friend masses) were respectively 476, 634, 167, 898, 80, 811, 45, 836 and 27, 955. For the correlation function we have shown only a simple Poissonian errors:

$$\Delta\xi \simeq \frac{1 + \xi}{\sqrt{DD}}. \quad (2.17)$$



**Figure 2.3.:** Upper panel: various corrections applied to reach the final power spectrum estimate. Lower panel: results of the consistency test for the “sharpening” scheme using different grid sizes.



**Figure 2.4.:** Upper panel: Power spectrum of clusters more massive than  $1.0 \cdot 10^{14} h^{-1} M_{\odot}$  from  $z = 0$  simulation box. Points with errorbars show the power spectrum of clusters inside a spherical subvolume with comoving radius of  $1500 h^{-1} \text{ Mpc}$  extracted from the full simulation box. Stars present results obtained by Colberg et al. (2000) and solid/dashed lines are model power spectra with/without acoustic oscillations. Lower panel: as above, except all the curves have been divided by the “smooth” model without acoustic oscillations.

(Errors due to the terms DR and RR can be neglected because of the much larger number of available pairs.) These errorbars are an underestimate of the true variance, being only exact for the Poissonian point process. To estimate true errors, one needs to have a knowledge about the 3- and 4-point correlation function of the clustering pattern. We are not going to elaborate further on these issues since the following analysis is based solely on the power spectrum. From Fig. 2.5 we see that baryonic features are visible down to the case with the lowest number of clusters. The only exception is the correlation function with the lowest number of objects which shows rather noisy behavior near the expected baryonic bump. For the least massive systems, due to the nonlinear evolution, the power spectrum starts to rise at  $k \gtrsim 0.15 h \text{ Mpc}^{-1}$ . For the more massive systems, on the other hand, the opposite trend is visible i.e. a decrease of power. This is due to the cluster formation which can be viewed as a smoothing filter acting on an initial density perturbation field.

### 2.2.2. Comparison with the analytical description: accuracy of the biasing scheme

The analytical power spectrum of clusters with masses  $M > M_{\text{low}}$  on a light-cone  $P_c^{\text{LC}}(k; > M_{\text{low}})$  is calculated as presented in Yamamoto et al. (1999) with a slight modification to allow for various weight functions, so:

$$P_c^{\text{LC}}(k; > M_{\text{low}}) = \frac{\int_{z_{\min}}^{z_{\max}} W^2(z) P_c(k; > M_{\text{low}}, z) \frac{dV_c}{dz} dz}{\int_{z_{\min}}^{z_{\max}} W^2(z) \frac{dV_c}{dz} dz}, \quad (2.18)$$

where the weight function  $W(z)$  is given in Eq.(2.8). There the number density of objects is provided by the cumulative mass function at redshift  $z$ :

$$n(> M_{\text{low}}, z) = \int_{M_{\text{low}}}^{\infty} n(M, z) dM. \quad (2.19)$$

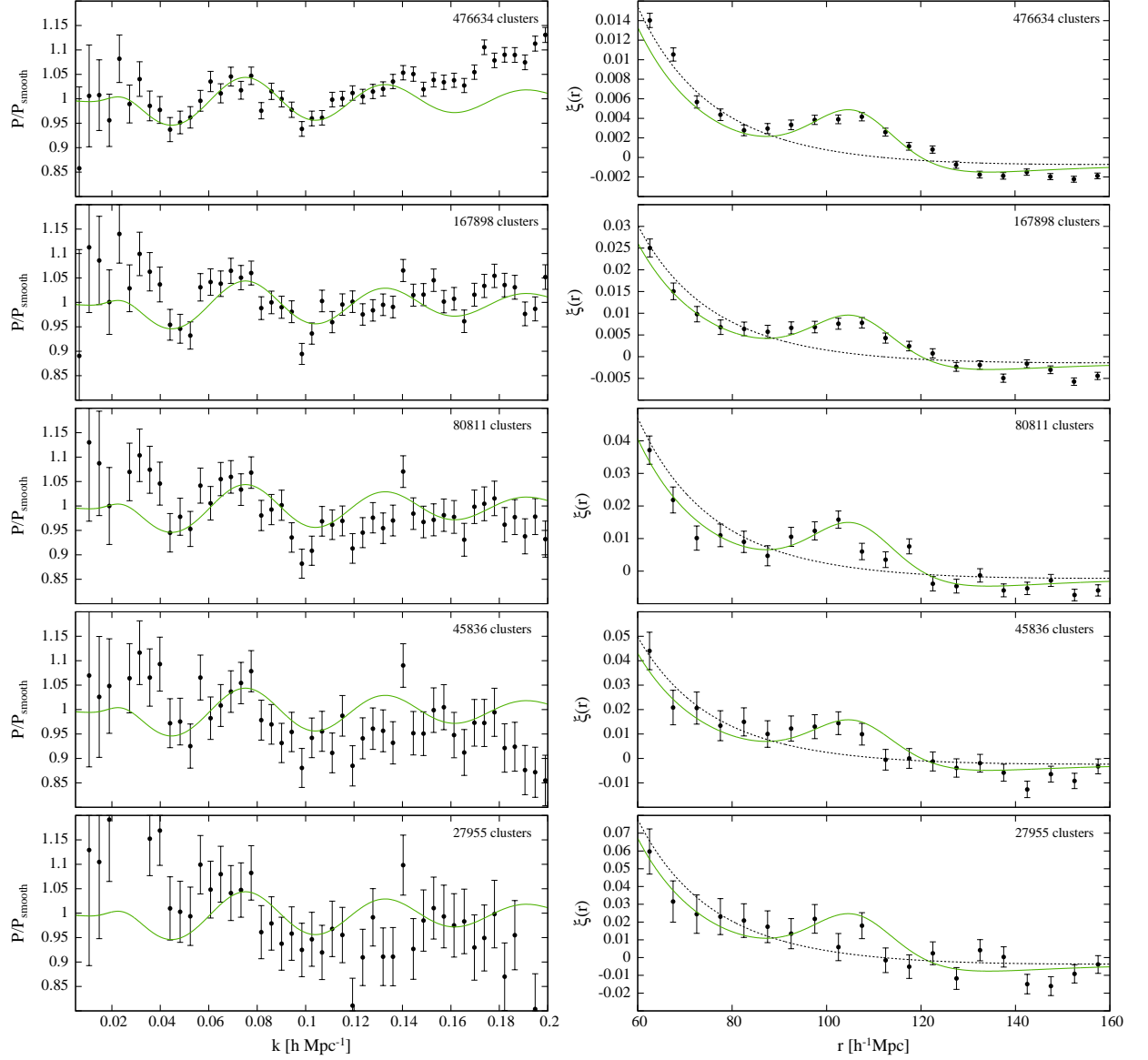
The power spectrum of clusters more massive than  $M_{\text{low}}$  at redshift  $z$  is given as:

$$P_c(k; > M_{\text{low}}, z) = D_+^2(z) b_{\text{eff}}^2(> M_{\text{low}}, z) P(k, z = 0), \quad (2.20)$$

where the effective bias parameter:

$$b_{\text{eff}}(> M_{\text{low}}, z) = \frac{\int_{M_{\text{low}}}^{\infty} b(M, z) n(M, z) dM}{n(> M_{\text{low}}, z)}. \quad (2.21)$$

$D_+(z)$  is the growing mode of linear density fluctuations normalized such that  $D_+(z = 0) = 1$  and  $P(k, z = 0)$  is the matter power spectrum at the current epoch, which is calculated using the transfer functions presented in Eisenstein & Hu (1998).



**Figure 2.5.:** Power spectra (left panels) and correlation functions (right panels) for the lower mass cutoffs of  $1.0 \cdot 10^{14}$ ,  $2.0 \cdot 10^{14}$ ,  $3.0 \cdot 10^{14}$ ,  $4.0 \cdot 10^{14}$  and  $5.0 \cdot 10^{14} h^{-1} M_{\odot}$  (friend-of-friend masses). Power spectra have been divided by the model spectra without acoustic features. The number of objects inside a spherical survey volume reaching  $z = 0.58$  is given in each panel. Solid/dotted lines show theoretical models with/without baryonic oscillations.

For the mass function  $n(M, z)$  and bias parameter  $b(M, z)$  we use both Press-Schechter (PS) (Press & Schechter 1974) and Sheth-Tormen (ST) (Sheth & Tormen 1999) prescriptions. It is well known that PS mass function underpredicts the number density of massive objects (Sheth & Tormen 1999). PS overestimates while ST underestimates the bias parameter for massive halos (especially at larger redshifts) (Sheth & Tormen 1999). PS underestimation of number density turns out to be approximately compensated for by its overestimation of bias parameter, and as such, we get the best agreement with the numerical light-cone power spectra using a plain PS approach. This is demonstrated in Fig. 2.6 where we show the light-cone power spectra for various values of lower mass cutoff  $M_{\text{low}}$ . Results in the upper panel apply for the SphereB cluster catalog (reaching redshift  $z \sim 0.58$ ) whereas the ones on the lower part of the figure correspond to the OctantB catalog (reaching  $z \sim 1.46$ ) of VIRGO Hubble Volume simulation outputs. We obtained the best agreement if clusters were selected using  $M_{180}$ , but in bias calculations virial mass  $M_{\text{vir}}$  was used instead.

Overall the agreement between the numerical results and an analytical description is better than 20%.

## 2.3. SZ clusters and baryonic oscillations

The “concordance” cosmological model predicts oscillations in the matter power spectrum with a relative amplitude of  $\sim 5\%$  (see Fig. 2.7). The correspondence of the peaks in the matter power spectrum to the ones in the CMB angular power spectrum is also given in Fig. 2.7 (The first vertical line represents the position of the 1st CMB acoustic peak etc.). We see that on small scales the corresponding features are out of phase. This is due to the so-called velocity overshoot, meaning that at those scales the growing mode of density fluctuations is mostly sourced by the velocity perturbations (Sunyaev & Zeldovich 1970, Eisenstein & Hu 1998). At larger scales, on the other hand, fluctuations in the density provide a dominant source term and so the corresponding features in the power spectra are in phase. Each rise and fall in the matter power spectrum corresponds to twice as many features in the CMB angular spectrum. This is a generic property of the models with a dominating CDM component. Purely baryonic models on the other hand would have oscillations with the same frequency as in the angular spectrum of the CMB.

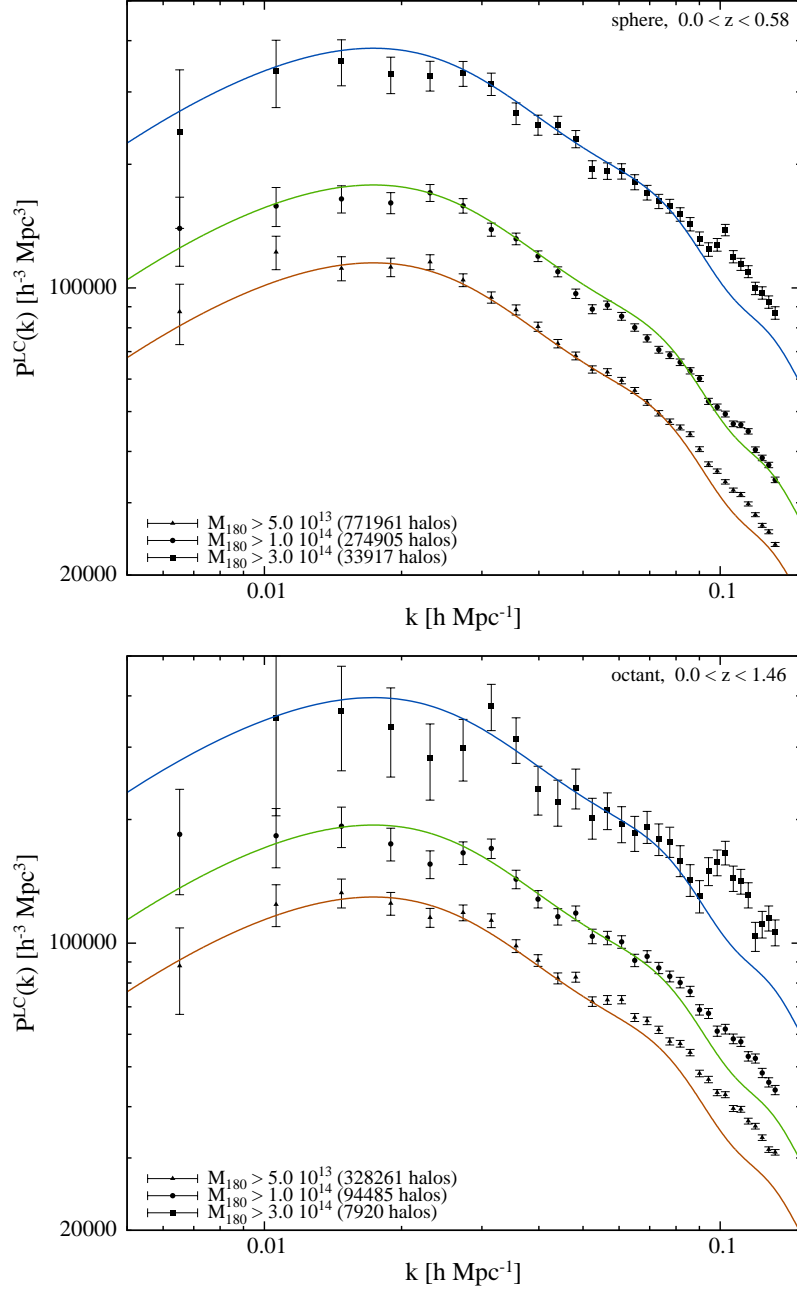
<sup>18</sup> In order to see these features in the matter power spectrum one needs a “tracer” population of objects whose clustering properties with respect to the underlying dark matter distribution are reasonably well understood. These objects should have high enough number density to reduce discreteness noise on one hand, but on the other hand they should fill as large of a comoving volume as possible to decrease cosmic variance. In general one wants to find an optimal solution of these two degrees of freedom <sup>19</sup>, as to maximize the obtainable effective volume (see Eq. (2.7)) for a fixed observational effort. Currently the largest effective volume amongst all of the available surveys is provided by the SDSS LRG sample (Eisenstein et al. 2005). The analysis of this sample yielded a detection of a clear acoustic feature in the spatial two point correlation function. Future projects such as the K.A.O.S.<sup>20</sup> galaxy redshift survey has as one of its main

<sup>18</sup><http://cmb.as.arizona.edu/~eisenste/acousticpeak/>

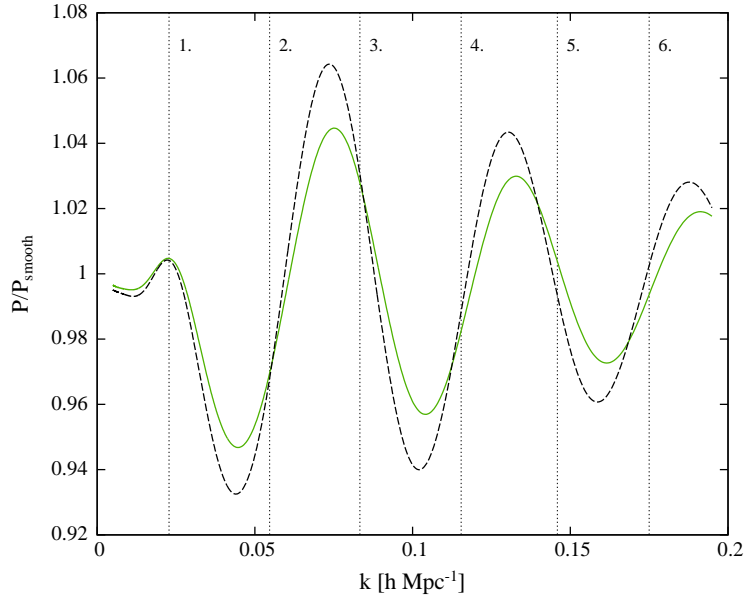
<sup>19</sup>One also has to consider the steepness of the luminosity function of those objects.

<sup>20</sup><http://www.noao.edu/kaos/>





**Figure 2.6.:** Upper panel: light-cone power spectra for SphereB output of VIRGO  $\Lambda$ CDM Hubble Volume simulation for different lower mass cutoffs (see legend). Lower panel: same as above but for OctantB output.



**Figure 2.7.:** Theoretical dark matter power spectra with a smooth component divided out (using transfer functions as given by Eisenstein & Hu (1998)) for the WMAP “concordance” model (solid line) and for the model used in VIRGO Hubble Volume simulations (dashed line). The numbered vertical lines show the locations of the corresponding peaks in the CMB angular power spectrum e.g. 1. corresponds to the 1st acoustic peak etc.

scientific targets the detection of baryon oscillations in the spatial clustering of high- $z$  galaxies. The possibility of using the aforementioned galaxy redshift surveys to measure the sound horizon has been discussed in several papers e.g. Blake & Glazebrook (2003), Linder (2003), Seo & Eisenstein (2003), Hu & Haiman (2003) and a similar discussion in the context of photometric redshift surveys is given in Blake & Bridle (2005) (see also the discussion in Seo & Eisenstein 2003).

Here instead of galaxies we discuss the possibility of using SZ-selected galaxy clusters for that purpose. Some calculations related to the SPT-type of SZ survey were also presented in Hu & Haiman (2003). It is clear from Figs. 2.4 and 2.7 that with wide field galaxy cluster surveys we should be especially sensitive to the scales that correspond to the 2nd and 3rd acoustic peaks in the CMB angular power spectrum.

A few advantages of using galaxy clusters compared to the galaxies are:

- With relatively small cluster samples it is possible to probe large cosmological volumes (thus reducing cosmic variance).
- The clustering signal of galaxy clusters is amplified with respect that of galaxies.
- The relation with respect to the underlying dark matter field is rather well understood and also redshift space distortions are manageable since “fingers of god” could be avoided.<sup>21</sup>

The biggest disadvantage is a rather low number density i.e. high shot noise contribution.

<sup>21</sup>This is only true when we have enough galaxy redshifts per cluster, so that one can average down to something

### 2.3.1. SZ-selected clusters. Mass-observable relations

In order to compare observations with the models one has to establish mass-observable relations and also specify survey selection criteria. Here for the sake of simplicity we assume that all the clusters remain unresolved i.e. we assume that our sample is effectively flux-selected. This is a rather good approximation for the case of PLANCK, but for surveys like ACT and SPT extra complications will arise since part of the cluster population will be resolved and so the selection function has one additional degree of freedom, namely surface brightness. As our aim here is not to give any detailed predictions for a particular survey these assumptions seem to be quite reasonable. The change in detected flux towards a galaxy cluster due to the thermal SZ effect can be expressed as:

$$F(M, z, x) = \frac{I_0 \sigma_T}{\mu_e m_p m_e c^2} \cdot \frac{g(x) f_b M \cdot kT(M, z)}{d_A(z)^2}, \quad (2.22)$$

where  $I_0 = 2(kT_{\text{cmb}})^3/(hc)^2 \simeq 2.7 \cdot 10^{11} \text{ mJy/sr}$ ,  $f_b$  is the cluster baryonic fraction which we take to be equal to the cosmic average  $\Omega_b/\Omega_m$ ,  $d_A(z)$  is angular diameter distance to the cluster,  $\mu_e = 2/(1+X)$  for the case of fully ionized plasma with negligible metallicity (we take  $\mu_e = 1.14$ ). The spectral function  $g(x)$  is given as follows (Sunyaev & Zeldovich 1980):

$$g(x) = \left[ x \coth\left(\frac{x}{2}\right) - 4 \right] \cdot \frac{x^4 e^x}{(e^x - 1)^2}, \quad (2.23)$$

where the dimensionless frequency  $x = h\nu/kT_{\text{cmb}} \simeq 0.0176 \cdot \nu(\text{GHz})$ . For the mass-temperature relation we assume a simple virial scaling (Bryan & Norman 1998):

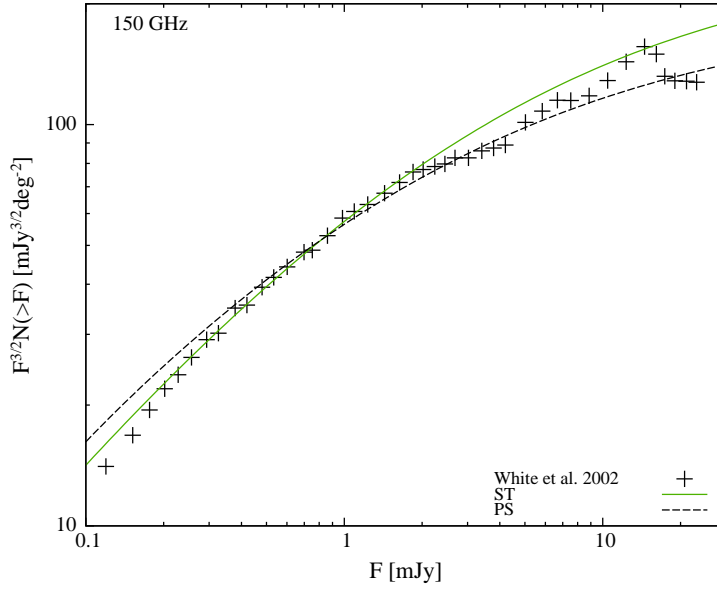
$$kT(M, z) = A \cdot \left[ \Delta_c(z) E(z)^2 \right]^{\frac{1}{3}} M^{\frac{2}{3}}. \quad (2.24)$$

Here  $\Delta_c(z)$  is a critical collapse overdensity with respect to the critical density at redshift  $z$  and  $E(z) = H(z)/H_0$ . The normalizing constant  $A$  is determined so as to obtain a good match for the SZ cluster number counts from the state-of-the-art hydrodynamical simulations by White et al. (2002). These simulations included gas cooling processes and also feedback from supernovae and galactic winds. If we measure  $kT$  in keV and  $M$  in units of  $h^{-1} M_\odot$ , then a good fit can be obtained if  $A \simeq 1.0 \cdot 10^{-10}$  as seen in Fig. 2.8. Here we present results both for ST and PS mass functions.

In reality the mass-observable relations are currently rather poorly known but one may argue that planned surveys with the yields of tens of thousands of galaxy clusters have significant power for “self-calibration” (Majumdar & Mohr 2004). As also shown in Majumdar & Mohr (2004) a much better approach would be to establish these scaling laws using external mass determinations (e.g. through lensing studies) for a subset of a complete sample. For a more precise modeling of the selection effects one also has to consider scatter around these mean relations. These issues can be settled once we obtain a real sample. Moreover, the clustering as

---

close to the center of mass velocity. One might also try to exploit the fact that the bright central cD galaxies have small velocities with respect to the rest of the cluster (see the simulation results by Berlind et al. (2003)). Even if one is able to find a good estimate for the center of mass velocity for each cluster in the sample, the redshift space distortions on reasonably large scales would still deviate from the simple linear prediction (Scoccimarro 2004).



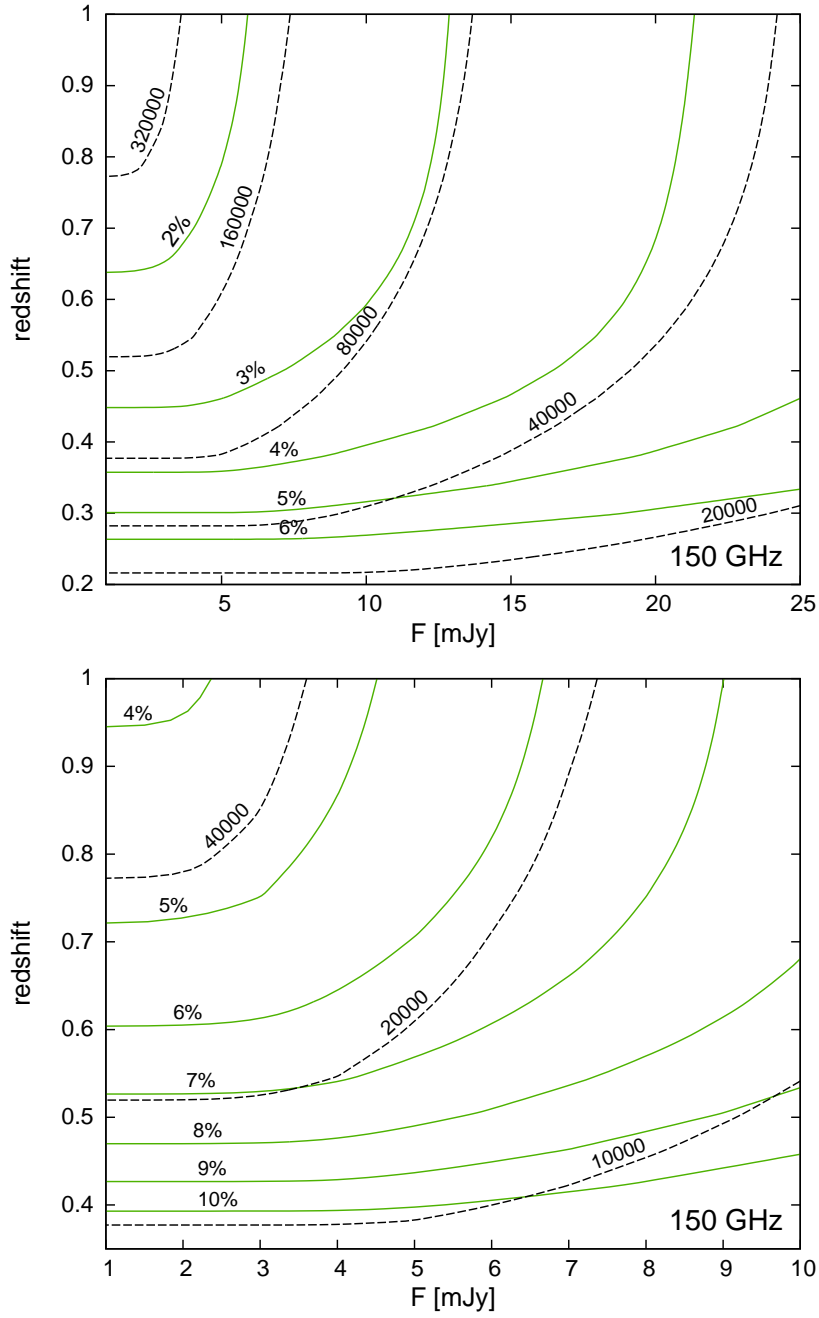
**Figure 2.8.:** Cumulative number counts of SZ clusters for observing frequency 150 GHz. Crosses show simulation results by [White et al. \(2002\)](#) whereas solid/dashed lines correspond to analytical results assuming a ST/PS mass function.

compared to the number count of objects is much less sensitive to the uncertainties in the precise knowledge of the selection effects. Here the selection effects enter when relating the clustering of tracer objects to the underlying dark matter i.e. while determining the effective bias of objects. As it turns out (see Sec. 4), future large cluster samples are able to provide a good estimate of the effective bias themselves through the redshift space distortions.

### 2.3.2. Accuracy of the power spectrum determination

We investigate how well we can determine the power spectrum with SZ surveys having various sensitivity limits, specifically concentrating on the range  $1 \dots 25$  mJy. Our results for  $\Delta P/P$  (or equivalently for the effective volume, see Eq. (2.6)) and for the number of detectable clusters are given in Fig. 2.9. Here the upper panel assumes full sky coverage and the lower one applies for one octant of the sky. With the solid lines we have plotted the fractional accuracy  $\Delta P/P$  achieved for different lower flux limits and follow-up survey depths applying the FKP weighting scheme. The wavenumber  $k$  in the calculations was taken to be  $0.05 h \text{ Mpc}^{-1}$ , which is close to the first major acoustic feature in the expected matter power spectrum (see Fig. 2.7). The bin width  $\Delta k = 0.005 h \text{ Mpc}^{-1}$  is large enough so that even for the shallowest surveys reaching only  $z \sim 0.2$  the power spectrum bins can be assumed to be independent. Here and also in the following we perform analytical calculations for the observational frequency 150 GHz and assume cosmological parameters consistent with the WMAP “concordance” model ([Spergel et al. 2003](#)).

Light-cone power spectra for SZ clusters are calculated using Eq. (2.18) where the lower integration boundary  $M_{\text{low}}(F, z)$  is given by Eq. (2.22) and (2.24). Then  $\Delta P/P$  is found using Eq. (2.6), (2.7) and (2.8) and the mean underlying number density  $\bar{n}(z)$  is given by Eq. (2.19). Also



**Figure 2.9.:** Upper panel:  $\Delta P/P$  (given in %) and number of clusters for flux-limited surveys with various sensitivities and survey depths at 150 GHz assuming full sky coverage. Solid lines show  $\Delta P/P$  isocontours whereas constant number contours are given with dashed lines. Lower panel: the same as above only for one octant of the sky.

we have taken into account the increase of the isotropized power spectrum due to linear redshift space distortions by a factor of  $1 + 2\beta/3 + \beta^2/5$  (Kaiser 1987), where  $\beta = \frac{1}{b_{\text{eff}}} \cdot \frac{d \ln D_+}{d \ln a}$ .

The solid lines in Fig. 2.9 starting from below correspond to  $\Delta P/P$  values of 6%, 5%, 4%, ... in the upper panel and 10%, 9%, 8%, ... in the lower one. With the dashed lines we have plotted the number of clusters. Moving from the lower right to the upper left each line represents a factor of two increase in number with the starting values being 20,000 and 10,000 in upper and lower panels, respectively. The flattening out of  $\Delta P/P$  and cluster number curves at low fluxes is due to the imposed lower mass cutoff  $1.0 \cdot 10^{14} h^{-1} M_{\odot}$ . Thus below some flux limit we see all the clusters inside a specified volume that have masses above that cutoff value.

In general, by increasing the volume of the survey we also boost the shot noise contribution due to the decreasing number density of distant objects. As seen from the figure- in the case of FKP weighting- the accuracy of power spectrum estimate does not degrade as we move to further distances since we downweight the contribution of the far away objects in such a way as to compensate for the increase in shot noise. For too small survey volumes, on the other hand, limits on achievable accuracy are set by the growing importance of the cosmic variance. The FKP weighting scheme is not strictly optimal in our case since it was derived assuming a fixed i.e. non-evolving underlying power spectrum. Certainly, for a better scheme one should weight down the contribution of the far away objects slightly more mildly since the clustering strength of these objects is higher. This type of weighting method, which is able to handle at least the case with an evolving amplitude, is presented in Percival et al. (2004). Nevertheless in the following calculations for simplicity we still apply the FKP weight function.

The results in Fig. 2.9 assumed that we have a full follow-up such that we are able to obtain all the redshifts of the clusters detected by an imaging survey. Also the sample was assumed to be purely flux-selected which is rather unrealistic for real experiments e.g. for PLANCK many clusters remain undetected due to rather poor angular resolution or oppositely in the case of SPT some fraction of clusters will be “resolved out” and a significant amount of signal will be lost. Many objects might remain undetected for these reasons. If the systems that are left out are low mass clusters (as in the case of PLANCK) then our power spectrum estimate might actually be almost as good as before since low mass systems are relatively weakly clustered and as such they do not contribute significantly to the total signal. This can be seen in Fig. 2.10 where in the upper panel we have shown the influence of changing the lower mass cutoff  $M_{\text{low}}$  on  $\Delta P/P$ . On the lower panel the respective number of clusters is given. These calculations are done for two different survey types: (1) solid lines represent results for a shallow SZ survey covering the full sky and reaching redshift  $z_{\text{lim}} = 0.6$  with a sensitivity limit  $F_{\text{low}} = 17$  mJy at 150 GHz, (2) dashed lines represent a deep survey (with no upper redshift cutoff imposed) covering 1/8 of the sky with a flux limit  $F_{\text{low}} = 5$  mJy at 150 GHz. The first might be applicable for the case of the PLANCK mission<sup>22</sup> and the second for the SPT cluster survey. In practice the measurements will be performed in many frequency channels which helps to separate clusters from other foreground sources due to the specific frequency behavior of the thermal SZ effect. Here for simplicity we have chosen the sensitivity limits corresponding to the “weakest” of the channels available for SZ purposes. We see that for a full sky with the 20,000 most massive clusters up to redshift  $z \sim 0.6$ , one could obtain an estimate of the power spectrum at  $k = 0.05 h \text{ Mpc}^{-1}$  with a fractional error

<sup>22</sup>Using the spectral dependence of thermal SZ effect (see Eq. (2.23)) we see that the 30 mJy sensitivity of PLANCK’s 353 GHz channel corresponds to 17 mJy at 153 GHz.

below 5% while for the one octant of the sky with less than  $\sim 25,000$  clusters we always stay above 5% accuracy.

The previously described approach where we use all the data to obtain a single combined estimate of the power spectrum is appropriate if we only intend to place constraints on  $\Omega_{\text{DE}}$ . Combining this power spectrum estimate which is sensitive to  $\Omega_{\text{m}}h$  with the CMB constraint on  $\Omega_{\text{m}}h^2$  gives us  $\Omega_{\text{m}}$  and  $h$  separately. Additionally, knowing the geometry of the Universe from CMB measurements gives us immediately an estimate of  $\Omega_{\text{DE}}$ . Since in the majority of the DE models DE starts to dominate relatively recently, driving the Universe furthest from the plain Einstein-de Sitter behavior, the best redshift to complement the CMB data is at  $z = 0$ .

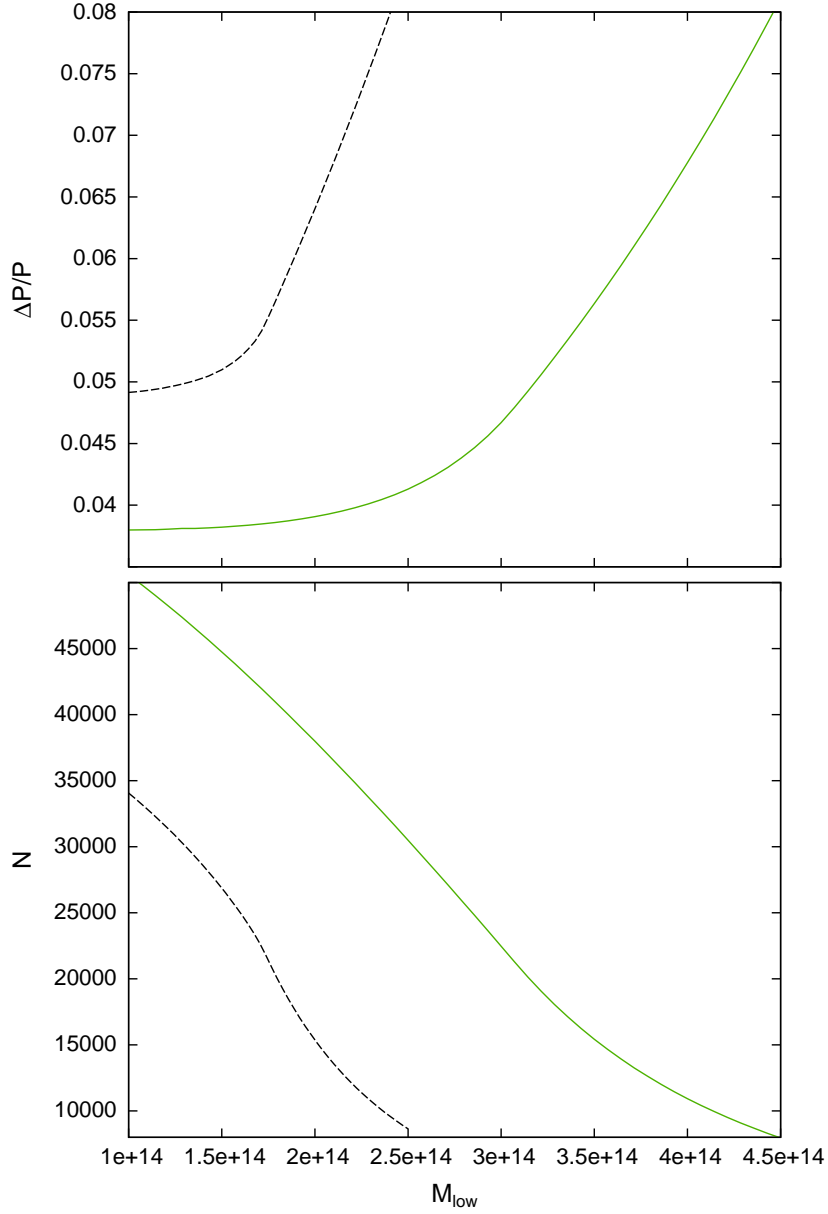
On the other hand if our aim is to constrain the equation of state parameter  $w$  and its possible change in time it is essential to measure the power spectrum at different redshifts. This leads to the question of how to bin up the sample in redshift? Certainly there are optimal ways of combining data, but unfortunately they all depend on the way we choose to parametrize our model for DE. Recently [Huterer & Starkman \(2003\)](#) argued that in the absence of a theoretically well motivated parametrization one should use a stepwise function with the value  $w_i$  in the  $i$ -th redshift bin and let the data itself determine which combinations of  $w_i$  will be well constrained. One can then reconstruct the behavior of  $w$  with the redshift as a linear combination of the “cleanest” eigenmodes. Here we are not trying to implement that kind of general parametrization since as a first step it should be sufficient to determine an effective constant  $w$  and see whether it deviates from the currently most well motivated  $w = -1$ . Therefore, in the following we mostly investigate the case with a constant equation of state parameter  $w_0$ . The redshift binning is chosen so as to get equal relative accuracies of the power spectrum in each bin. The results of this binning procedure for the above described two types of survey are given in Fig. 2.11 with the solid lines corresponding to the shallow one. For the shallow (deep) survey we have assumed 3 (4) redshift bins. In the inset the upper curves show the relative accuracy achievable in each redshift bin while the lower lines correspond to the full sample without any binning and so coincide with the lines shown in the upper panel of Fig. 2.10. This kind of redshift division is also used in the following parameter estimation section.

### 2.3.3. Prospects of detecting baryonic “wiggles”. Comparison with SDSS LRG

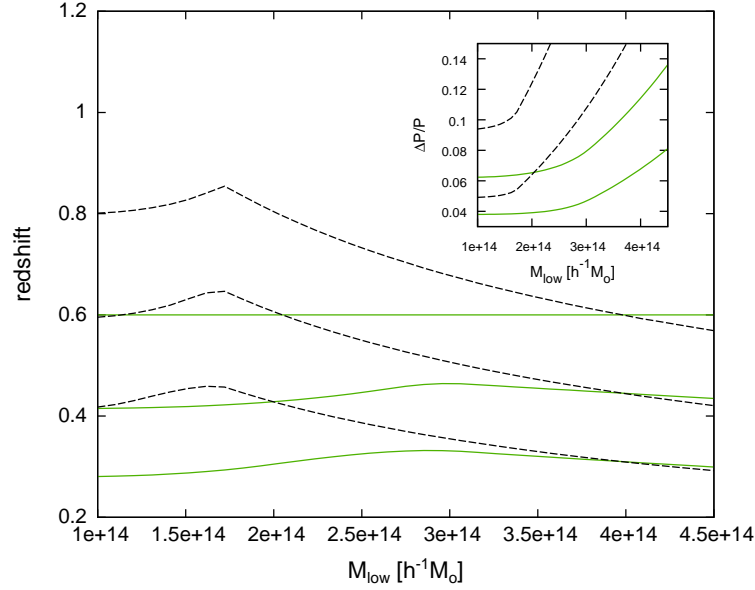
The relative amplitude of acoustic oscillations in the mater power spectrum for the WMAP “concordance” model as well as in the power spectrum used in VIRGO Hubble Volume simulations is shown in Fig. 2.7. Here the smooth component was divided out using the fitting formulae provided by [Eisenstein & Hu \(1998\)](#). In order to be able to detect these features the accuracy of the power spectrum determination should be of comparable size i.e.  $\Delta P/P \sim 5\%$ . This implies that for a full sky survey one needs on the order of 25,000 galaxy clusters inside the volume with limiting redshift  $z \sim 0.6$  as seen from the upper panel of Fig. 2.10. It is clear that with clusters one might hope to detect only the first few acoustic signatures, e.g. the major features at  $k \sim 0.045$  and  $\sim 0.075 h \text{ Mpc}^{-1}$ , since they are too rare objects to enable the sampling of the smaller scale density field.

Since the study of the SDSS LRG sample has led to the detection of acoustic oscillations in





**Figure 2.10.:**  $\Delta P/P$  (upper panel) and number of clusters (lower panel) as a function of the lower mass cutoff  $M_{\text{low}}$ . Solid lines correspond to the full sky shallow SZ survey and dashed lines to the deep survey covering one octant of the sky.

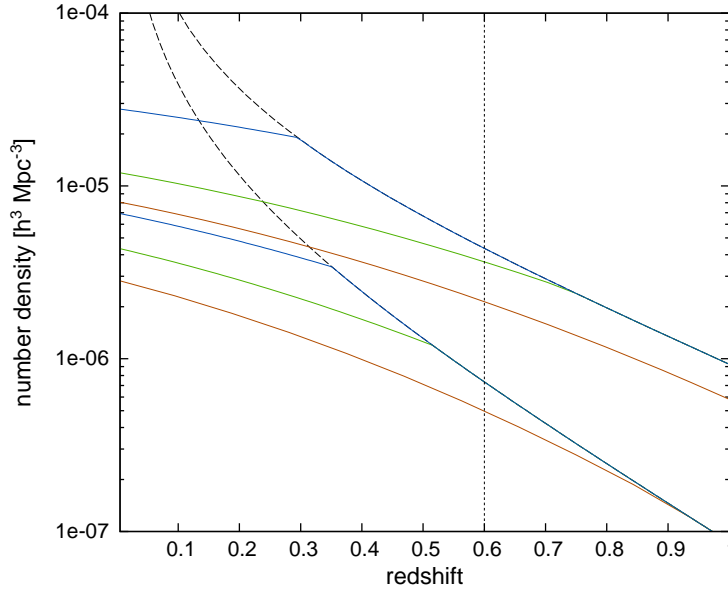


**Figure 2.11.:** “Equal accuracy” redshift binning for different lower mass cutoffs. Solid (dashed) lines correspond to the shallow (deep) SZ survey with 3 (4) redshift bins. In the inset the upper curves show the relative accuracy achievable in each redshift bin while the lower lines correspond to the full sample without any binning.

the spatial distribution of galaxies (Eisenstein et al. 2005),<sup>23</sup> it would be instructive to compare the “strength” of this survey to the planned blank sky SZ cluster surveys like PLANCK and SPT. In Fig. 2.12 we show the number density of clusters as a function of redshift for the above described two types of SZ survey. The upper group of lines corresponds to the SPT-like deep survey while the lower curves are for a wide and shallow survey like PLANCK. For each of the surveys we have varied the lower mass cutoff  $M_{\text{low}}$  so as to obtain in total 15,000, 25,000 and 35,000 clusters. These three cases are shown with solid lines. Dashed lines display the pure flux-limited surveys without any lower cutoff in the mass imposed. Using these results and also taking into account the proper biasing factors as given by the square root of Eq. (2.18) divided by  $P(k, z = 0)$ <sup>24</sup> we can readily obtain  $\Delta P/P$  as given by Eq. (2.6) and (2.7). The results of this calculation are given in Fig. 2.13. Here the solid lines correspond to the PLANCK-like and dashed ones to the SPT-like surveys. Each set of lines corresponds to the detected cluster numbers (starting from above): 15,000, 25,000 and 35,000. The dash-dotted curve, showing the results for the SDSS LRG sample, is found using Eq. (2.6) and the data for the effective volume given in Fig. 1 of Eisenstein et al. (2005) (again  $\Delta k = 0.005 h \text{ Mpc}^{-1}$  was assumed). We can see that on large scales future SZ surveys have enough strength to improve the results obtained using the SDSS LRG sample. However, the SDSS LRG sample is going to double in size within a few years as the survey is completed. The achievable  $\Delta P/P$  for this final sample is shown in Fig. 2.13 as a dotted line. Moreover, it seems that acoustic oscillations are able to survive at

<sup>23</sup>The detection has also been claimed using the 2dF redshift survey (Cole et al. 2005).

<sup>24</sup>The light-cone bias parameters calculated this way are 5.0, 4.7, 4.3 for the PLANCK and 4.4, 4.1, 3.9 for the SPT with the number of clusters 15,000, 25,000 and 35,000, respectively.

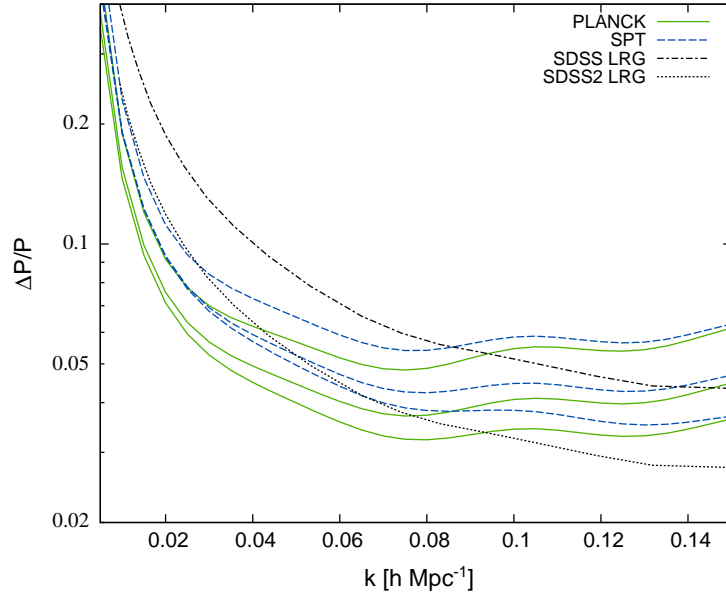


**Figure 2.12.:** Comoving number density of clusters as a function of redshift for the PLANCK- (lower set of curves) and SPT-like (upper set of curves) SZ surveys. Solid lines in each set correspond to the cases with 15,000, 25,000 and 35,000 clusters. The vertical dotted line shows the applied upper limiting redshift for the PLANCK. Dashed lines represent purely flux-limited cases i.e. without any lower mass cutoff imposed.

the quasilinear scales ( $k \sim 0.1 \dots 0.3 h \text{ Mpc}^{-1}$ ) (see Fig. 2.4) which significantly increases the amount of information available for the galaxy redshift surveys. This is also confirmed by the recent N-body simulations by [Springel et al. \(2005\)](#) and [Seo & Eisenstein \(2005\)](#). In order to fully exploit this information one needs a complete theoretical understanding of how nonlinear effects, redshift space distortions and nonlinear biasing influence these features. So far there have been only a few works studying these important issues (e.g. [Meiksin et al. 1999](#), [Springel et al. 2005](#), [Seo & Eisenstein 2005](#), [White 2005](#)) and we do not have a full theoretical description of them available yet. For this reason we have not attempted to incorporate the SDSS LRG sample into our Fisher matrix parameter estimation process.

### 2.3.4. Some remarks on SZ vs. optical cluster selection

SZ cluster selection might be superior to the simple optical one since SZ brightness does not suffer from ordinary cosmological dimming. This allows one to obtain an approximately mass-limited sample of clusters that is spatially much more uniform than optically selected samples. For example, in Fig. 2.12 the number density of clusters for both PLANCK (up to  $z \sim 0.6$ ) and the SPT type of surveys drops approximately as  $\propto z^{-2.5}$ . In contrast, the number density of the optically selected SDSS LRG sample (which contains mostly galaxies that populate dense cluster environments) drops as  $\propto z^{-5}$  beyond  $z \sim 0.3$ . So in general with SZ-selected clusters one is able to probe larger volumes. Also as a larger part of the sample is at higher redshifts the clustering signal is stronger due to the increase of the bias factor with increasing distance.



**Figure 2.13.:** Comparison of the “strength” of the future SZ cluster surveys with respect to the SDSS LRG. Lines shown for the PLANCK- and SPT-like surveys correspond to the detected cluster numbers (starting from above): 15,000, 25,000 and 35,000.

(Of course the question of whether these far away parts of the sample have high enough number density in order to be useful for the clustering study depends on the specific parameters of the experiment.) The other weaknesses of the optical selection are projection effects and confusion with the background objects. As an example, comparison of the X-ray and optically selected cluster catalogs often yields a rather poor match (e.g. [Donahue et al. 2002](#)), which probably signals that many of the “optically constructed” systems are actually false detections.

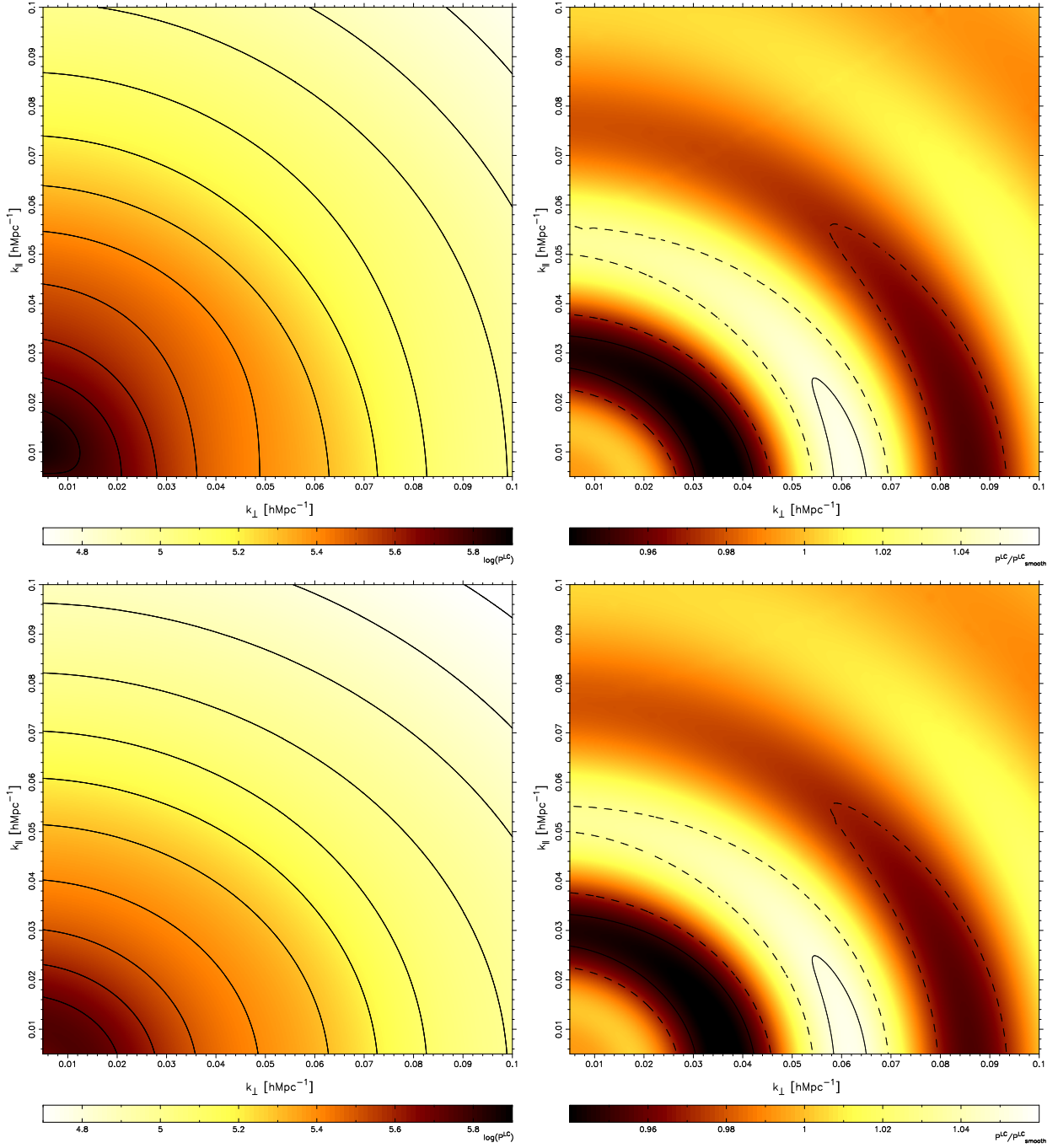
For the SZ surveys the spectroscopic follow-up is a crucial issue, but several other studies (such as cluster number counts) also require determination of redshifts. As these investigations will be performed anyway, the clustering study can be seen as coming essentially “for free”.

## 2.4. Constraints on Dark Energy

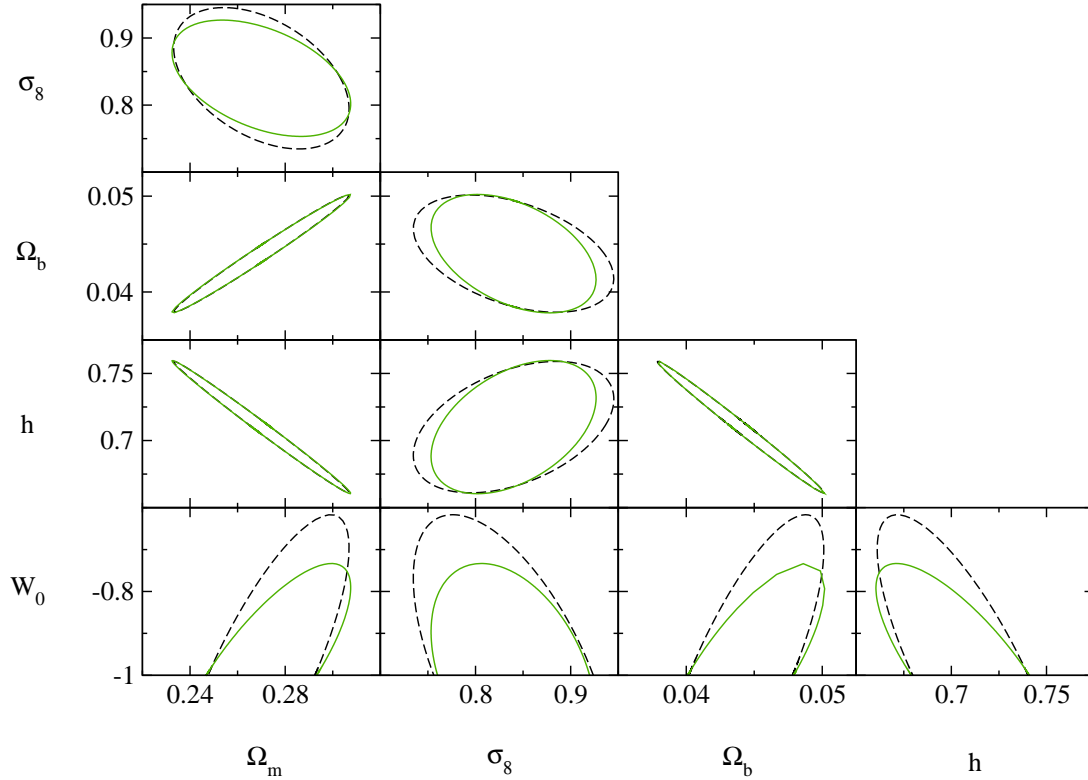
### 2.4.1. 2D power spectrum on a light-cone

Having obtained a parametrized (and well calibrated) analytical model for the light-cone power spectra of SZ-selected clusters of galaxies, we can estimate the accuracy with which it is possible to recover cosmological parameters. Since the observations are done in cosmological (as opposed to the comoving) redshift space there are two additional effects one has to take into account:

1. The increase of power along the line of sight due to the large scale coherent inflows towards massive accretion centers. This effect is accounted for using results from linear theory.



**Figure 2.14.:** Upper left panel: 2D light-cone power spectrum for clusters with mass above  $1.0 \cdot 10^{14} h^{-1} M_{\odot}$  in logarithmic units. Lower left panel: as above, but the redshift distortion “switched off”. Upper right panel: as upper left panel, only smooth component of the spectrum divided out and results shown using linear scale. Lower right panel: the same procedure applied to the lower left panel.



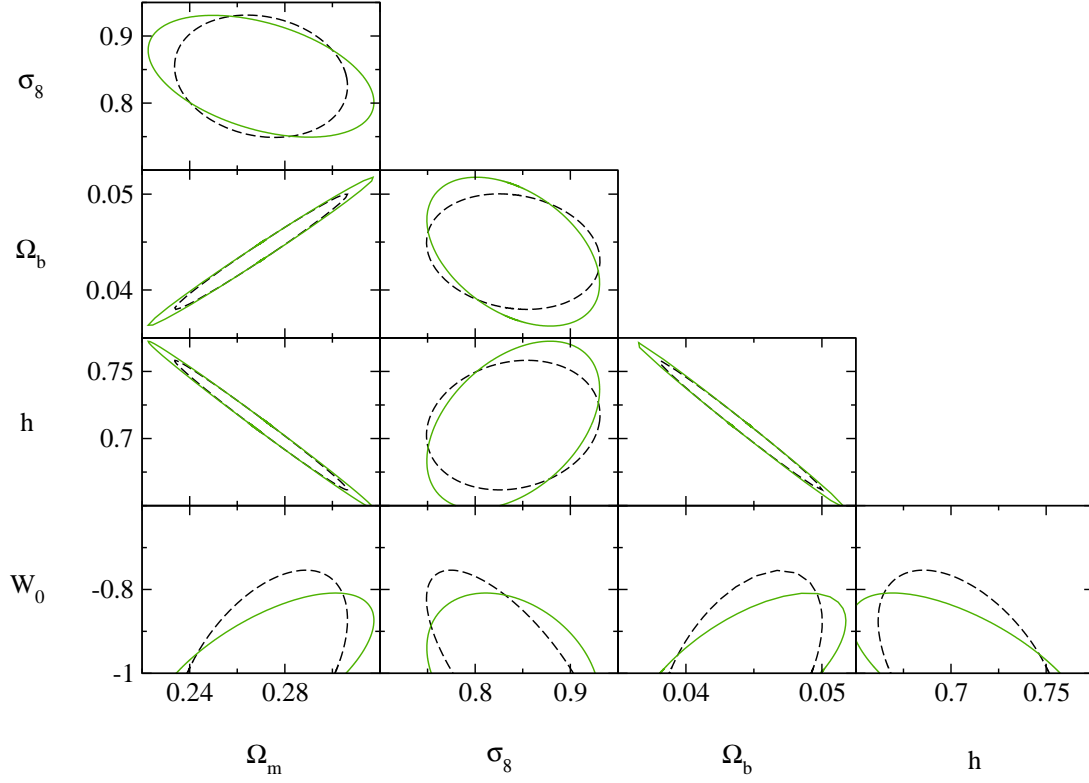
**Figure 2.15.:** Error ellipses for the PLANCK-like SZ survey including CMB priors as described in the text. Solid/dashed lines show results for 3/1 redshift bin(s).

2. Cosmological distortion due to the fact that one can directly observe only redshifts, and in order to find the corresponding comoving distances, one has to assume some cosmological model. Choosing an incorrect model will lead to distortions along and perpendicular to the line of sight <sup>25</sup>.

Both of these effects will, in general, lead to an anisotropic power spectrum. Thus instead of a one dimensional (isotropized) power spectrum one has to consider here a two dimensional power spectrum with components along and perpendicular to the line of sight. For simplicity, in the following we use a flat sky approximation. The 2D power spectra are calculated following the description given in [Magira et al. \(2000\)](#) with a slight modification to allow for various weight functions <sup>26</sup>. [Magira et al. \(2000\)](#) took the fiducial cosmology given by the Milne's empty universe model. Thus the comoving distance intervals along and perpendicular to the line

<sup>25</sup>The Alcock-Paczynski test ([Alcock & Paczynski 1979](#)) is based on these cosmological distortions.

<sup>26</sup>In the calculations we use the FKP weight function.



**Figure 2.16.:** Analog of Fig 2.15 for the SPT-like survey. Solid/dashed lines show results for 4/1 redshift bin(s).

of sight are:

$$\Delta x_{\parallel}^{\text{ref}} = \frac{c}{H_0} \Delta z, \quad (2.25)$$

$$\Delta x_{\perp}^{\text{ref}} = \frac{c}{H_0} z \Delta \theta, \quad (2.26)$$

where  $\Delta z$  is redshift interval and  $\Delta \theta$  angular separation between two objects. For the general FRW universe the corresponding intervals read as:

$$\Delta x_{\parallel} = \frac{c}{H(z)} \Delta z, \quad (2.27)$$

$$\Delta x_{\perp} = d_M(z) \Delta \theta, \quad (2.28)$$

where  $d_M(z) = (1+z)d_A(z)$  is the comoving transverse distance. Now, defining the shift parameters:

$$c_{\parallel}(z) = \frac{\Delta x_{\parallel}}{\Delta x_{\parallel}^{\text{ref}}}(z), \quad (2.29)$$



$$c_{\perp}(z) = \frac{\Delta x_{\perp}}{\Delta x_{\perp}^{\text{ref}}}(z), \quad (2.30)$$

we can write down the final expression for the 2D power spectrum on a light-cone:

$$P^{\text{LC,2D}}(k_{\parallel}, k_{\perp}; > M_{\text{low}}) = \frac{\int_{z_{\min}}^{z_{\max}} dz \frac{dV_c}{dz} W^2(z) \left[ 1 + \beta(> M_{\text{low}}, z) \left( \frac{k'_{\parallel}}{k'} \right)^2 \right]^2 \cdot P_c(k'; > M_{\text{low}}, z)}{\int_{z_{\min}}^{z_{\max}} dz \frac{dV_c}{dz} W^2(z) c_{\perp}(z)^2 c_{\parallel}(z)}, \quad (2.31)$$

where

$$k'_{\parallel} = \frac{k_{\parallel}}{c_{\parallel}(z)}, \quad k'_{\perp} = \frac{k_{\perp}}{c_{\perp}(z)}, \quad k' = \sqrt{k'^2_{\parallel} + k'^2_{\perp}} \quad (2.32)$$

and

$$\beta(> M_{\text{low}}, z) = -\frac{1}{b_{\text{eff}}(> M_{\text{low}}, z)} \cdot \frac{d \ln D_+(z)}{d \ln(1+z)}. \quad (2.33)$$

The factor  $c_{\perp}(z)^2 c_{\parallel}(z)$  in the denominator of Eq. (2.31) is the Jacobian determinant taking into account the change in a volume element. It is missing in the numerator due to the cancellation by the similar but inverse term arising from the transformation of the  $k$ -space volume element. The term in square brackets models the amplification due to the coherent inflows and the last term in the numerator,  $P_c(k'; > M_{\text{low}}, z)$ , is given earlier by Eq. (2.20). In the case of SZ flux-selected clusters,  $M_{\text{low}}$  at each redshift for a given lower flux limit is found using Eq.(2.22) and (2.24).

In Fig. 2.14 we present some examples of 2D power spectra calculated in the manner described above. The top left-hand panel shows the 2D light-cone power spectrum of clusters with a mass above  $1.0 \cdot 10^{14} h^{-1} M_{\odot}$  (in logarithmic units) while the lower left-hand panel contains the same spectrum but with the redshift space distortion “switched off”. The contours starting from the upper right corner correspond to the values  $10^{4.8}$  and  $10^{4.7}$  for the upper and lower panel, respectively and the step size was taken  $10^{0.1} h^{-3} \text{Mpc}^3$ . It is clearly seen how linear redshift-space distortion boosts power along the line of sight. The cosmological distortion (in the currently selected reference model) on the other hand works in the opposite way. Since the chosen Milne model is strongly different from the  $\Lambda$ CDM “concordance” model the cosmological distortion is easily visible. We use Milne’s model only for illustrative purposes but in the following parameter estimation part we change the reference model to the WMAP “concordance” cosmology. On the right-hand panels of the figure we have removed the smooth component of the power spectrum revealing the series of acoustic rings. Here the continuous contours correspond to  $P^{\text{LC,2D}}/P^{\text{LC,2D}}_{\text{smooth}}$  values of 0.95 and 1.05 while the dashed lines are for the values 0.97 and 1.03. It is important to note that the picture on both panels looks practically the same. This is due to the fact that cosmological and redshift-space distortions work in a different way: cosmological transformation stretches or compresses the power spectra on the plane of the figure whereas redshift distortion moves the spectra in a vertical direction. If we had smooth power spectra without any particular features then it would be extremely hard to disentangle these two types of distortions. Having the power spectra with acoustic features it is easy to isolate cosmological distortion by dividing out a smooth component. The ability to disentangle cosmological and redshift-space distortions is extremely important to extract the bias parameter from the survey in a self-consistent way. In the parameter estimation part of this section we see how much better one does with the model having acoustic oscillations as compared to the one without.

The total power spectrum measured over a broad  $z$ -interval is a weighted sum of differently distorted power spectra and so there will be some loss of acoustic features. The loss is stronger along the line of sight, as can be seen from Fig. 2.14. This is due to the currently chosen reference model where  $|\frac{dc_{\parallel}(z)}{dz}| > |\frac{dc_{\perp}(z)}{dz}|$ . Again, this effect is strongly pronounced because Milne's model differs strongly from the WMAP “concordance” cosmology for which our calculation was done.

## 2.4.2. Parameter estimation

In this subsection we apply a Fisher matrix forecasting techniques to study how well one can determine cosmological parameters. Since the pioneering investigations in the field of CMB anisotropies (Jungman et al. 1996) and galaxy redshift surveys (Tegmark 1997) these methods have gained great popularity. For a full description of the method with applications see Tegmark et al. (1997). Following Tegmark (1997) the Fisher matrix in the case of a 2D power spectrum can be written as (see also Hu & Haiman (2003)):

$$F_{ij} = \sum_n \sum_m \frac{\partial \ln P^{\text{LC},2\text{D}}(k_{\parallel n}, k_{\perp m})}{\partial \Theta_i} \cdot \frac{V_k^m V_{\text{eff}}}{2} \cdot \frac{\partial \ln P^{\text{LC},2\text{D}}(k_{\parallel n}, k_{\perp m})}{\partial \Theta_j}, \quad (2.34)$$

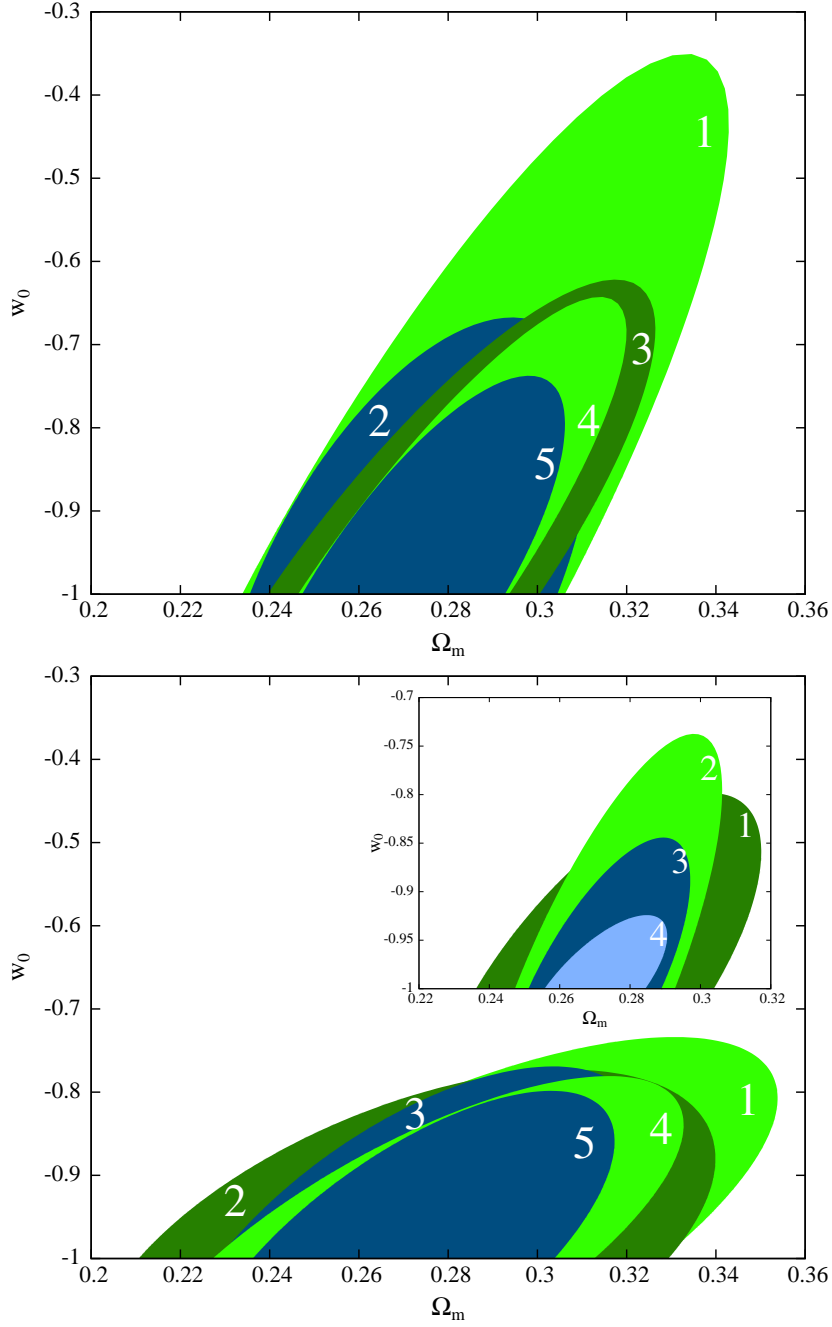
where

$$V_k^n = \frac{4\pi k_{\perp n} \Delta k_{\perp} \Delta k_{\parallel}}{(2\pi)^3}, \quad (2.35)$$

and we take  $\Delta k_{\perp} = \Delta k_{\parallel} = \Delta k$ .  $P^{\text{LC},2\text{D}}$  is given in Eq. (2.31) and  $V_{\text{eff}}$  in Eq. (2.7). The wavevector components  $k_{\parallel i}$  and  $k_{\perp j}$  form a rectangular grid with a step size  $\Delta k$ . We allow them to span the range  $0.005 \dots 0.1 h \text{ Mpc}^{-1}$  and take  $\Delta k = 0.005 h \text{ Mpc}^{-1}$  i.e. we have a  $20 \times 20$  grid. It is not justified to go to higher wavenumbers than  $0.1 h \text{ Mpc}^{-1}$  since there already the simple linear scaling of the cluster power spectrum seems to break down (see Fig. 2.6). The parameters  $\Theta_i$  are taken to be  $\Omega_m$ ,  $\sigma_8$ ,  $\Omega_b$ ,  $h$ ,  $w_0$  plus the bias parameters  $b_n$  in each redshift bin. Thus in the case of four redshift bins (which is the maximum number considered in our analysis) we have a total of nine free parameters. We also assume a fiducial cosmology given by the best-fit WMAP “concordance” model (Spergel et al. 2003) plus the dark energy in the form of the cosmological constant i.e.  $w = -1$ .

Here we perform calculations for two types of SZ survey. The first one that should serve as a prototype for PLANCK is a shallow survey that covers the full sky with a sensitivity limit of  $F_{\text{low}} = 30 \text{ mJy}$  at  $353 \text{ GHz}$ <sup>27</sup>. We also apply a lower mass cutoff of  $M_{\text{low}} = 2.75 \cdot 10^{14} h^{-1} M_{\odot}$  resulting in total  $\sim 25,000$  clusters up to the applied limiting redshift  $z \simeq 0.6$  (see Fig. 2.10 lower panel). The calculations are done for one redshift bin spanning  $z = 0 \dots 0.58$  and also for three bins:  $z_1 = 0 \dots 0.33$ ,  $z_2 = 0.33 \dots 0.45$ ,  $z_3 = 0.45 \dots 0.58$ . The achieved  $\Delta P/P$  values at  $k = 0.05 h \text{ Mpc}^{-1}$  are 4.3% for the one bin case and 7.5% for each of the bins in the three bin case. The other deep and rather “narrow” survey should mimic the performance of SPT. Here we have taken  $F_{\text{low}} = 5 \text{ mJy}$  at  $150 \text{ GHz}$  and  $M_{\text{low}} = 1.5 \cdot 10^{14} h^{-1} M_{\odot}$  (see Ruhl et al. (2004), Majumdar & Mohr (2004)). Moreover, we assumed that the survey is capable of covering one octant of the sky and here we have not applied any high redshift cutoff. As for the PLANCK-like survey two different cases are considered here. In the first case while combining all the data to

<sup>27</sup><http://astro.estec.esa.nl/Planck>



**Figure 2.17.:** Upper panel: Error ellipses under various assumptions for the PLANCK-like survey. Here we have shown the following cases: 1. Clustering signal of the model without baryonic oscillations + CMB priors, 2. #1 + prior on bias, 3. Clustering signal of the “wiggly” model, 4. #3 + CMB priors, 5. #4 + prior on bias. Lower panel: as above, except for the SPT-like survey and with a changed order of the cases #2 and #3. Lower panel inset: constraints obtainable with: 1. SPT, 2. PLANCK, 3. SPT + PLANCK, 4. The survey with SPT characteristics covering the full sky. All of the cases here have CMB and bias priors added.

**Table 2.1.:** Principal components as given in Eq. (2.36) for the “clustering only” case for PLANCK-like, SPT-like and for the combined survey.

	i	$\lambda^i$	$e_1^i$	$e_2^i$	$e_3^i$	$e_4^i$	$e_5^i$
PLANCK (3 $z$ -bins)	1	$1.6 \cdot 10^3$	0.74	-0.0064	-0.24	0.62	0.090
	2	$1.7 \cdot 10^2$	-0.48	-0.015	-0.43	0.49	-0.58
	3	32.	0.38	-0.53	-0.23	-0.46	-0.56
	4	18.	0.027	0.41	-0.80	-0.37	0.26
	5	5.9	0.27	0.74	0.28	-0.13	-0.53
SPT (4 $z$ -bins)	1	$9.5 \cdot 10^2$	0.71	-0.094	-0.23	0.63	0.21
	2	85.	0.15	-0.088	0.29	-0.36	0.87
	3	54.	-0.012	-0.99	-0.032	-0.10	-0.13
	4	17.	0.69	0.094	0.24	-0.54	-0.41
	5	4.6	0.049	0.057	-0.90	-0.42	0.12
PLANCK+ SPT	1	$2.6 \cdot 10^3$	0.73	-0.038	-0.24	0.63	0.13
	2	$2.5 \cdot 10^2$	-0.38	0.044	-0.39	0.45	-0.70
	3	78.	0.18	-0.92	-0.027	-0.22	-0.28
	4	52.	0.54	0.39	0.087	-0.44	-0.60
	5	24.	0.021	0.057	-0.89	-0.40	0.22

one common power spectrum we are able to achieve  $\Delta P/P \simeq 5.1\%$ , whereas in the second case with the four redshift bins:  $z_1 = 0 \dots 0.45$ ,  $z_2 = 0.45 \dots 0.63$ ,  $z_3 = 0.63 \dots 0.83$ ,  $z_4 = 0.83 \dots$  we find a  $\Delta P/P \simeq 9.9\%$  in each redshift bin. The total number of objects detected with this type of survey would be  $\sim 27,000$ .

The results in Fig. 2.15 and 2.16 present constraints for the five cosmological parameters:  $\Omega_m$ ,  $\sigma_8$ ,  $\Omega_b$ ,  $h$  and  $w_0$ . Here we have added priors to  $\Omega_m h^2$  and  $\Omega_b h^2$  from the CMB angular power spectrum studies, and also a prior to the bias parameters. In order to be able to compare our results to the ones given in Hu & Haiman (2003) the fractional errors of 0.01 for  $\Omega_m h^2$  and  $\Omega_b h^2$  (this should be achievable with the PLANCK mission (Hu 2002)) were similarly assumed. These constraints are easily “rotated” to the frame used here since under coordinate transformations the Fisher matrix transforms as a second rank tensor. Moreover, we have restricted our calculations to the flat models only, and for the bias have assumed that one is able to describe it with a relative accuracy of 15%. Finally, the joint Fisher matrix is the sum of all the Fisher matrices transformed to a common frame. All the calculations done here assume an underlying model with baryonic features in the matter power spectrum. Fig. 2.15 presents results for the previously described shallow survey with full sky coverage while the error ellipses in Fig. 2.16 apply to the deep and narrow survey. In both figures dashed lines correspond to the single and solid lines to the multiple bin case. Due to the fact that with a single bin one is able to measure the shape of the power spectrum with a higher precision than in the case of multiple bins, we see from the above figures that in general stronger constraints on  $\Omega_m$  are obtained. On the other hand constraints on  $w_0$  are much stronger in the multiple bin case due to the increased knowledge about the redshift derivatives. We also performed calculations taking two redshift bins and the results were already rather close to the solid curves in the figures above which can be interpreted as an indication that

any further redshift slicing would not improve constraints on  $w_0$ . Also one should not increase the number of redshift bins much above the maximally used values of three and four since then the wavevector bins with width  $\Delta k = 0.005 h \text{ Mpc}^{-1}$  would become highly correlated and the above Fisher matrix calculation would not be meaningful.

Probably the most interesting constraints in Fig. 2.15 and 2.16 are the ones for  $\Omega_m$  (or  $\Omega_{DE}$  since we have assumed flat models) and  $w_0$ . Error ellipses in the  $\Omega_m$ - $w_0$  plane are also given in Fig. 2.17 for various different assumptions. The top panel here corresponds to the PLANCK-type and the lower one to the SPT-type of survey. The order of ellipses in the top panel starting from the bottommost one is as follows: (1) clustering signal of the model without baryonic oscillations + CMB priors, (2) (1) + prior on bias, (3) clustering signal of the “wiggly” model, (4) (3) + CMB priors, (5) (4) + prior on bias. The only difference in the lower panel is the reversed order of (2) and (3). The inset in the lower plot displays the constraints obtainable (again starting from the bottommost ellipse) with (1) SPT, (2) PLANCK, (3) SPT + PLANCK, (4) the survey with SPT characteristics covering the full sky. For all of the cases shown in the inset we have included CMB and bias priors. The constraint ellipses for the “clustering only” case assuming a “smooth” model would fill almost all the plot area and for the sake of clarity we have not displayed them here. It is evident from Fig. 2.17 that the model with baryonic oscillations is performing much better compared to its smoothed counterpart. Adding prior information (in contrast to the “smooth” case) does not result here in a strong improvement i.e. the clustering signal alone already has a significant constraining power.

Table 2.1 lists the principal components of the clustering analysis only i.e. no CMB and bias priors included. Also we have marginalized over bin bias parameters. This serves as a compact way of summarizing our results. The principal components are given in the form:

$$\prod_{j=1}^5 \left( \frac{\Theta_j}{\Theta_j^{\text{fid}}} \right)^{e_j^i} = 1 \pm \frac{1}{\sqrt{\lambda^i}} \quad \text{for all } i = 1 \dots 5. \quad (2.36)$$

Here  $e_j^i$  is the  $j$ -th component of the  $i$ -th eigenvector and  $\lambda^i$  is the  $i$ -th eigenvalue of the matrix  $\widetilde{F}_{ij} = \Theta_i^{\text{fid}} \Theta_j^{\text{fid}} F_{ij}$  (no summation over indices). The parameter vectors  $\Theta = (\Omega_m, \sigma_8, \Omega_b, h, w_0)$  and  $\Theta^{\text{fid}} = (0.27, 0.84, 0.044, 0.71, -1.0)$ .

#### 2.4.2.1. Comparison to previous work

Probably the two closest works to ours are Majumdar & Mohr (2004) and Hu & Haiman (2003). In Majumdar & Mohr (2004) the authors discuss constraints obtainable by combining the cluster power spectrum with independent information from cluster number counts. Unfortunately they do not present results separately for the clustering signal only. Moreover, they use an isotropized power spectrum which leads to a significant loss of information, especially when the spectra contain baryonic features. In Hu & Haiman (2003) the authors use the full 2D power spectrum of galaxy clusters although they do not take into account light-cone effects. Since the light-cone power spectrum is a blend of differently deformed power spectra (if we are away from the reference model point) some loss of baryonic features will result as seen from Fig. 2.14. Also in their analysis they used  $k$ -modes up to 0.15 as opposed to our adopted value of  $0.1 h \text{ Mpc}^{-1}$ . As seen from Fig. 2.6 at  $k = 0.15 h \text{ Mpc}^{-1}$  the cluster power spectrum already differs quite significantly from the simple linear one, particularly for the more massive systems. To simplify the

comparison our results with [Hu & Haiman \(2003\)](#), we used identical CMB priors. An additional difference is that we do not allow the spectral index of the power spectrum to vary but keep it fixed to  $n = 1$ . In total their results for the SPT-type of survey are significantly more optimistic, e.g. the constraints on  $w_0$  agree roughly with our SPT+PLANCK case, however, in the case of  $\Omega_m$  an approximate agreement is achieved with our SPT full sky example.

## 2.5. Conclusions

In this paper we studied the clustering of SZ-selected galaxy clusters on a past light-cone with particular emphasis on constraining the properties of DE. We implemented an extended Press-Schechter type of analytical model as described e.g. in [Sheth & Tormen \(1999\)](#). The description of the calculation of the light-cone power spectra (e.g. [Yamamoto et al. 1999](#)) was modified slightly to incorporate other than simple number weighting schemes. The analytical model was extensively calibrated using the outputs from the VIRGO Consortium’s Hubble Volume simulations. With a little bit of fine tuning we were able to match analytical light-cone power spectra with ones extracted from the simulations to an accuracy better than 20%. The SZ scaling relations were calibrated so as to get a good match to the number count results from the state-of-the-art hydrodynamical simulations of [White et al. \(2002\)](#). Having a well calibrated analytical model we investigated how accurately future SZ surveys like PLANCK and SPT could determine the cluster power spectrum and whether they would be able to detect traces of baryonic oscillations. Also we made use of VIRGO simulation outputs to build cluster catalogs for various survey depths and sensitivity limits. We showed that the aforementioned blank sky SZ surveys will be able to improve the detection of acoustic features based on the SDSS LRG sample. To obtain a high-fidelity detection of the baryonic oscillations, one has to reach a relative accuracy  $\sim 5\%$  if the wavenumber bin width  $\Delta k = 0.005 h \text{ Mpc}^{-1}$  is used. This seems hard to achieve with surveys like SPT still having relatively narrow sky coverage. On the other hand, for surveys with a wide sky coverage the prospects seem to be better e.g. with  $\sim 25,000$  most massive clusters up to redshift  $\sim 0.6$  one should be able to reach a relative accuracy  $\sim 4.5\%$  at  $k = 0.05 h \text{ Mpc}^{-1}$  i.e. roughly the scale where one expects to see the first major acoustic feature.

In the last part of the paper we carried out a Fisher matrix forecasting analysis for cosmological parameters, which add up to nine in the case of four redshift bins:  $\Omega_m$ ,  $\sigma_8$ ,  $\Omega_b$ ,  $h$ ,  $w_0$  plus a free bias parameter for each bin. We included prior information for  $\Omega_m h^2$  and  $\Omega_b h^2$  from CMB angular power spectrum studies and also constrained the possible values for the bias parameters. A prior on bias parameters only has a significant effect in the case of models with smooth power spectra i.e. models with acoustic oscillations have enough constraining power to give an estimate of bias parameters from the survey itself. The most interesting constraints are obtained for  $\Omega_m$  and  $w_0$ . Wide and rather shallow surveys like PLANCK in combination with a CMB prior on  $\Omega_m h^2$  are able to provide strong constraints on  $\Omega_m$  or in the case of flat models equivalently on  $\Omega_{DE} = 1 - \Omega_m$ . The constraints on  $w_0$  on the other hand are not as good as the ones obtained by deeper and narrower surveys with the characteristics of SPT due to the lack of higher redshift objects. We also give our results as the principal components of the Fisher matrix that should allow for an easy way of comparison and also for a fast way of incorporating these constraints into further parameter forecasting studies.





# 3. Acoustic oscillations in the SDSS DR4 Luminous Red Galaxy sample power spectrum

G. Hütsi

*Astronomy & Astrophysics 449, 891 (2006)*

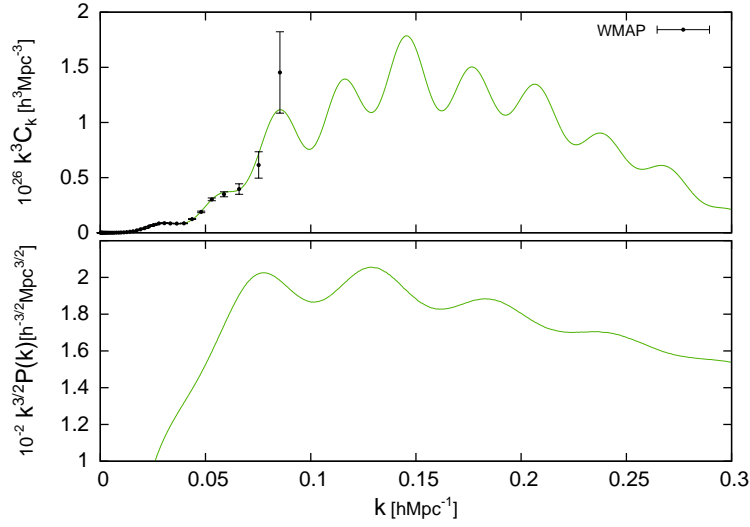
The analysis for the SDSS DR3 was carried out in [Hütsi \(2005\)](#).

## Abstract

We calculate the redshift-space power spectrum of the Sloan Digital Sky Survey (SDSS) Data Release 4 (DR4) Luminous Red Galaxy (LRG) sample, finding evidence for a full series of acoustic features down to the scales of  $\sim 0.2 h \text{ Mpc}^{-1}$ . This corresponds up to the 7th peak in the CMB angular power spectrum. The acoustic scale derived,  $(105.4 \pm 2.3) h^{-1} \text{ Mpc}$ , agrees very well with the “concordance” model prediction and also with the one determined via the analysis of the spatial two-point correlation function by [Eisenstein et al. \(2005\)](#). The models with baryonic features are favored by  $3.3\sigma$  over their “smoothed-out” counterparts without any oscillatory behavior. This is not only an independent confirmation of [Eisenstein et al. \(2005\)](#) results made with different methods and software but also, according to our knowledge, the first determination of the power spectrum of the SDSS LRG sample.

## 3.1. Introduction

In the beginning of 1970’s it was already realized that acoustic waves in the tightly coupled baryon-photon fluid prior to the epoch of recombination will lead to the characteristic maxima and minima in the post-recombination matter power spectrum. The same mechanism is also responsible for the prominent peak structure in the CMB angular power spectrum ([Sunyaev & Zeldovich 1970](#), [Peebles & Yu 1970](#), [Doroshkevich et al. 1978](#)). The scale of these features reflects the size of the sound horizon, which itself is fully determined given the physical densities  $\Omega_b h^2$  and  $\Omega_m h^2$ . The acoustic horizon can be calibrated using the CMB data, thus turning it into a standard ruler which can be used to carry out classical cosmological tests. For example, if we are able to measure the redshift and angular intervals corresponding to the physically known acoustic scale in the matter power spectrum at a range of redshifts, we can immediately find angular diameter distance  $d_A$  and Hubble parameter  $H$  as a function of redshift. Having good knowledge of these dependencies allows us to put constraints on the properties of the dark energy. To carry out this project one needs a tracer population of objects whose clustering properties with respect to the underlying matter distribution is reasonably well understood. There have been several



**Figure 3.1.:** Acoustic oscillations in the CMB (upper panel) and linear matter power spectrum (lower panel) for the “concordance” cosmological model. Here, as we have plotted the spectra against spatial wavenumber  $k$ , we have changed the standard notation of  $C_\ell$  to  $C_k$ . Due to the  $k^3$  factor the first CMB acoustic peak is barely visible. Density fluctuations in matter at smaller scales, being mostly induced by the velocity fields, are out of phase with respect to the fluctuations in the CMB component. Also the fluctuation period is twice as large.

works discussing the usage of galaxies (Blake & Glazebrook 2003, Hu & Haiman 2003, Linder 2003, Seo & Eisenstein 2003) and clusters of galaxies (Hu & Haiman 2003, Majumdar & Mohr 2004, Hütsi 2006b) for this purpose. What is most important is that already currently existing galaxy redshift surveys have lead to the detection of acoustic features in the spatial distribution of galaxies, this way providing clearest support for the standard gravitational instability picture of the cosmic structure formation. In the paper by Eisenstein et al. (2005) the detection of the acoustic “bump” in the two-point redshift-space correlation function of the SDSS <sup>1</sup> LRG sample is announced. The discovery of similar features in the power spectrum of 2dF <sup>2</sup> galaxies is presented in Cole et al. (2005). These results clearly demonstrate the great promise of the future dedicated galaxy redshift surveys like K.A.O.S. <sup>3</sup> Similarly, useful measurements of the acoustic scale can be hoped by the planned SZ cluster surveys like the ones carried out by the PLANCK Surveyor <sup>4</sup> spacecraft and SPT <sup>5</sup> (Hütsi 2006b) and also with a large future photometric redshift surveys (Blake & Bridle 2005). For the SZ surveys one needs an additional optical follow-up to get estimates for the cluster redshifts. In this paper we calculate the redshift-space power spectrum of the SDSS LRG sample finding evidence for the acoustic oscillations down to the scales of  $\sim 0.2 h \text{ Mpc}^{-1}$ , which effectively correspond up to the 7. peak in the CMB angular power spectrum. These scales in the CMB are very strongly damped due to the finite width of

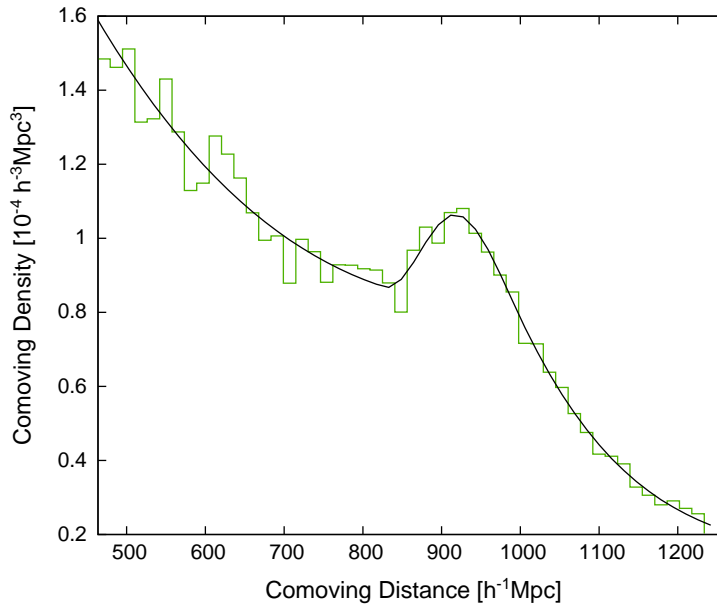
<sup>1</sup><http://www.sdss.org/>

<sup>2</sup><http://www.mso.anu.edu.au/2dFGRS/>

<sup>3</sup><http://www.noao.edu/kaos/>

<sup>4</sup><http://astro.estec.esa.nl/Planck>

<sup>5</sup><http://astro.uchicago.edu/spt>



**Figure 3.2.:** Comoving number density of galaxies as a function of comoving distance. Smooth solid line shows a cubic spline fit to the number density estimated for 50 discrete radial bins.

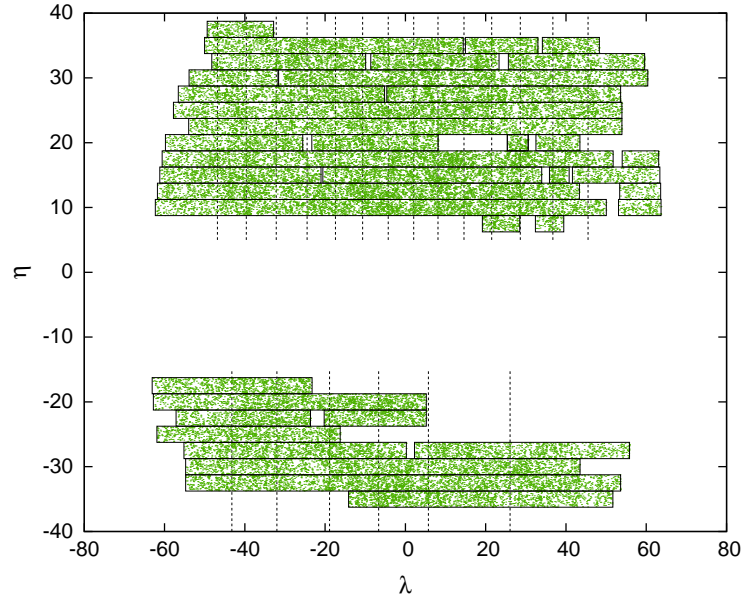
the last-scattering surface and also due to the Silk damping (Silk 1968). This can be seen in Fig. 3.1<sup>6</sup> where the CMB data is plotted in a somewhat unusual way to enhance the acoustic features at the high wavenumber damping tail. Also, at those scales the secondary CMB anisotropies (mostly thermal Sunyaev-Zeldovich effect (Sunyaev & Zeldovich 1972, 1980)) start to dominate over the primary signal. On the other hand, features in the matter power spectrum, although being small ( $\sim 5\%$  fluctuations), are preserved by the linear evolution and so opening up the way to probe acoustic phenomena at scales smaller than the ones accessible for the CMB studies.

The paper is structured as follows. In Sec. 2 we describe the dataset to be analyzed. Sec. 3 presents the method of the power spectrum calculation. In Sec. 4 we determine power spectrum errors and covariance matrix. Sec. 5 discusses the convolution effect of the survey window. Analytical model spectra are presented in Sec. 6. The results of the measurement of the acoustic scale are given in Sec. 7. Correlation function analysis is carried out in Sec. 8. In Sec. 9 we compare the measured power spectrum with the published results for the 2dF and SDSS main sample, and finally we conclude with Sec. 10.

## 3.2. Data

We analyze the publicly available data from the SDSS DR4 (Adelman-McCarthy et al. 2005). Specifically, we carry out our power spectrum measurements using the subset of the SDSS spectroscopic sample known as the Luminous Red Galaxy (LRG) sample. The LRG selection algo-

<sup>6</sup>Here instead of the usual multipole number  $\ell$  we have plotted the CMB angular power spectrum against the wavenumber  $k$ . For the “concordance” cosmological model  $\ell = 9990 k[h \text{ Mpc}^{-1}]$ .



**Figure 3.3.:** Angular distribution of galaxies given in the SDSS survey coordinates  $(\lambda, \eta)$ . The survey mask is shown with solid lines. The vertical dashed lines show the division of the sample into 22 separate regions each containing  $\sim 2350$  galaxies. This division will be exploited in the “jackknife” error analysis of the correlation function.

rithm (Eisenstein et al. 2001) selects  $\sim 12$  galaxies per square degree meeting specific color and magnitude criteria <sup>7</sup>. The resulting set of galaxies consists mostly of an early types populating dense cluster environments and as such are significantly biased (bias factor  $b \sim 2$ ) with respect to the underlying matter distribution. The selection method is very effective producing a galaxy sample with a reasonably high density up to the redshift of  $z \sim 0.5$ .

Since the selection criteria are very complicated, involving both cuts in magnitude and in color, and also due to the steepness of the luminosity function the usual method of using only the luminosity function to determine radial selection function does not work here (Zehavi et al. 2005). Here we simply build the radial selection function as a smooth spline fit to the number density profiles shown in Fig. 3.2. To calculate distances we choose the cosmological parameters as given by the WMAP <sup>8</sup> “concordance” model (Spergel et al. 2003). Unfortunately the coverage masks of the SDSS DR4 spectroscopic sample are not available in a readily accessible format and so we chose to build the angular survey masks using the galaxy data itself <sup>9</sup>. As the number density of galaxies in the sample is rather high, one can determine relatively accurately the beginning, ending and also possible gaps in the scan stripes. We have built angular masks using

<sup>7</sup>For the exact details of the selection criteria see Eisenstein et al. (2001)

<sup>8</sup><http://lambda.gsfc.nasa.gov/product/map/>

<sup>9</sup>In principle one can build the angular masks using the raw tiling information, but as we show later our approximate treatment is probably rather fine, since the results seem to be quite stable against small uncertainties in the mask. More rigorous approach should certainly address the issues of survey boundaries and completeness fluctuations (expected to be small due to the very effective tiling algorithm by Blanton et al. (2003)) in a much better detail.

both the whole DR4 galaxy sample and LRGs only. The measured power spectra are practically identical with only some minor differences on smaller scales (see Fig. 3.6). This can be seen as an indication that our power spectrum measurements are rather stable against small uncertainties in the survey geometry. The angular distribution of the galaxies and also the boundaries of the survey mask built in the above mentioned way (here using all the galaxies) is shown in Fig. 3.3. Here the angular positions are plotted using the so-called survey coordinate system of the SDSS <sup>10</sup>.

We have selected all the objects that have spectrum classified as galaxy (i.e. `SpecClass=2`) and are additionally flagged as `GALAXY_RED` or `GALAXY_RED_II` (i.e. `PrimTarget` bit mask set as `0x20` or `0x40000000`, respectively). Only galaxies for which the redshift confidence parameter, `zConf`, is greater than 0.95 were used. We apply lower and upper redshift cutoffs of 0.16 and 0.47 as also done in Eisenstein et al. (2005). The lower cutoff is needed since the color cuts that define the LRG sample break down for redshifts below  $\sim 0.2$  (Eisenstein et al. 2001). For the analysis presented in this paper we have excluded the three southern stripes since these just increase the sidelobes of the survey window without adding much of the extra volume. We have also removed some minor parts of the sample to obtain more continuous and smooth chunk of volume. In total the analyzed galaxy sample covers  $\sim 0.75 h^{-3} \text{ Gpc}^3$  over  $\sim 3850$  square degrees on the sky and contains 51,763 galaxies.

### 3.3. Power spectrum calculation

We calculate the power spectrum using a direct Fourier method as described in Feldman et al. (1994) (FKP). Strictly speaking, power spectra determined this way are the so-called pseudospectra, meaning that the estimates derived are convolved with a survey window. Since in the case of the analyzed LRG sample the volume covered is very large, reaching  $0.75 h^{-3} \text{ Gpc}^3$ , and also the survey volume has relatively large dimensions along all perpendicular directions, the correlations in the Fourier space are rather compact. On intermediate scales and in the case the power spectrum binning is chosen wide enough, FKP estimator gives a good approximation to the true underlying power.

The FKP estimate for a 3D pseudospectrum reads as:

$$\tilde{P}(\mathbf{k}) = |F(\mathbf{k})|^2 - P_{\text{shot}}, \quad (3.1)$$

where

$$F(\mathbf{k}) = \int d^3r F(\mathbf{r}) \exp(i\mathbf{k} \cdot \mathbf{r}). \quad (3.2)$$

Here  $F(\mathbf{r})$  is the weighted density contrast field:

$$F(\mathbf{r}) = w(\mathbf{r}) [n_g(\mathbf{r}) - \alpha n_s(\mathbf{r})]. \quad (3.3)$$

$n_g(\mathbf{r})$  and  $n_s(\mathbf{r})$  denote the number densities of the analyzed galaxy catalog and a synthetic random catalog with the same selection criteria, respectively. Since we are dealing with discrete

<sup>10</sup>The transformations between various coordinate systems used by the SDSS are given e.g. in Stoughton et al. (2002).

point processes, densities can be given as:

$$n_g(\mathbf{r}) = \sum_i \delta^D(\mathbf{r} - \mathbf{r}_i^g), \quad (3.4)$$

$$n_s(\mathbf{r}) = \sum_i \delta^D(\mathbf{r} - \mathbf{r}_i^s), \quad (3.5)$$

where  $\mathbf{r}_i^g$  and  $\mathbf{r}_i^s$  denote the location of the  $i$ -th point in real and synthetic catalog, respectively, and  $\delta^D$  is the 3D Dirac delta function.  $\alpha$  in Eq. (3.3) is the ratio of the number of galaxies to the number of random points in the synthetic catalog i.e.  $\alpha = \frac{N_g}{N_s}$ . In our calculations we have  $N_s = 10^7$  and thus  $\alpha \simeq 0.0052$ . For the weight function  $w(\mathbf{r})$  there have been traditionally three choices in the literature:

$$w(\mathbf{r}) \propto \begin{cases} \frac{1}{\bar{n}(\mathbf{r})} & \text{for volume weighting} \\ \text{const} & \text{for number weighting} \\ \frac{1}{1+\bar{n}(\mathbf{r})\bar{P}} & \text{for an optimal FKP weighting.} \end{cases} \quad (3.6)$$

Here  $\bar{n}(\mathbf{r})$  is the average number density of galaxies at comoving location  $\mathbf{r}$  i.e. the radial selection function of the survey (see Fig. 3.2) times the angular mask (Fig. 3.3). In our calculations we use an optimal FKP weighting scheme, although pure volume weighting would give practically the same results, especially on the larger scales ( $k \lesssim 0.09 h \text{ Mpc}^{-1}$ ), since then for the majority of the sample  $\bar{n}(\mathbf{r})\bar{P} \sim 3$ <sup>11</sup>. The weights in Eq. (3.3) are normalized such that:

$$\int d^3r \bar{n}(\mathbf{r}) w^2(\mathbf{r}) = 1, \quad (3.7)$$

which can be approximated as the following sum over the synthetic catalog<sup>12</sup>:

$$\alpha \sum_i \bar{n}(r_i^s) w^2(r_i^s) = 1. \quad (3.8)$$

The last term in Eq. (3.1) represents the Poissonian discreteness noise and can be expressed as:

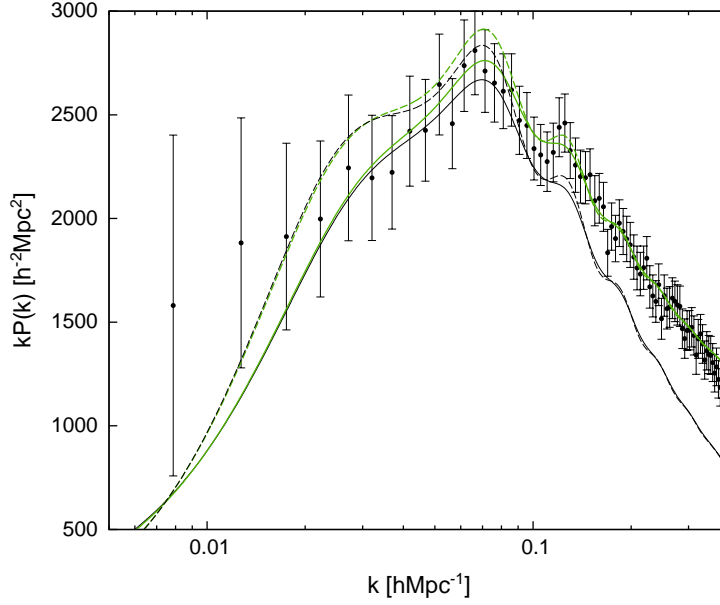
$$P_{\text{shot}} = (1 + \alpha) \int d^3r \bar{n}(\mathbf{r}) w^2(\mathbf{r}) \simeq \alpha(1 + \alpha) \sum_i w^2(r_i^s). \quad (3.9)$$

Since we are using Fast Fourier Transforms (FFTs) to speed up the calculation of the Fourier sums, we have to deal with some extra complications. As the density field is now “restricted to live” on a regular grid with a finite cell size, we have to correct for the smoothing effect this has caused. Also, if our underlying density field contains spatial modes with higher frequency than our grid’s Nyquist frequency,  $k_{\text{Ny}}$ , then these will be “folded back” into the frequency interval the grid can support, increasing power close to  $k_{\text{Ny}}$ —the so-called aliasing effect. The relation between the spectra calculated using direct summation and the ones found using FFT techniques was worked out by [Jing \(2005\)](#). It can be expressed as follows:

$$|F(\mathbf{k})|_{\text{FFT}}^2 = \sum_{\mathbf{n} \in \mathbb{Z}} |\mathcal{W}(\mathbf{k} + 2k_{\text{Ny}}\mathbf{n})|^2 \tilde{P}(\mathbf{k} + 2k_{\text{Ny}}\mathbf{n}) + P_{\text{shot}} \sum_{\mathbf{n} \in \mathbb{Z}} |\mathcal{W}(\mathbf{k} + 2k_{\text{Ny}}\mathbf{n})|^2, \quad (3.10)$$

<sup>11</sup>Including all the modes down to the scales of  $k \sim 0.25 h \text{ Mpc}^{-1}$  the effective value for  $\bar{n}(\mathbf{r})\bar{P}$  drops down to  $\sim 1.5$ .

<sup>12</sup>We assume that the survey selection does not have any other angular dependence except for the applied angular mask i.e. we can replace  $\mathbf{r}_i^s$  by the modulus  $r_i^s$ .



**Figure 3.4.:** Power spectrum of the SDSS LRG sample with the bin width  $\Delta k \approx 0.005 h \text{ Mpc}^{-1}$ . The upper solid line shows the best fitting model spectrum and the lower one corresponds to the linearly evolved matter power spectrum of the “concordance” cosmological model multiplied by the square of the bias parameter  $b = 1.95$ . Both of the spectra are convolved with a survey window. The dashed lines represent the corresponding unconvolved spectra.

where  $\mathcal{W}(\mathbf{k})$  is the mass assignment function used to build density grid out of the point set. We use the Triangular Shaped Cloud (TSC) assignment method (Hockney & Eastwood 1988). Since the TSC filter can be obtained by convolving uniform cube (the Nearest Grid Point filter) two times with itself, the Fourier representation of it follows immediately:

$$\mathcal{W}(\mathbf{k}) = \left[ \frac{\prod_{i=1}^3 \sin\left(\frac{\pi k_i}{2k_{Ny}}\right)}{\left(\frac{\pi k_i}{2k_{Ny}}\right)} \right]^3, \quad \mathbf{k} = (k_1, k_2, k_3). \quad (3.11)$$

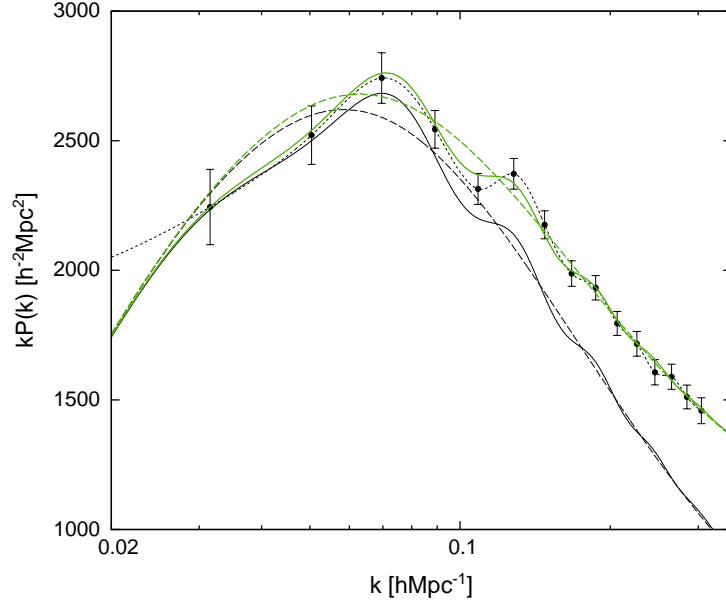
Here the sum that represents the contribution from aliases runs over all the integer vectors  $\mathbf{n}$ . Eq. (3.10) is the direct analog of the previous Eq. (3.1). The convolution with the mass assignment filter has introduced  $\mathcal{W}^2(\mathbf{k})$  factors both to the power spectrum and to the shot noise term. The sum in the last term of Eq. (3.10) can be performed analytically for the TSC filter to yield the result (Jing 2005)<sup>13</sup>:

$$\sum_{\mathbf{n} \in \mathbb{Z}} |\mathcal{W}(\mathbf{k} + 2k_{Ny}\mathbf{n})|^2 = \prod_{i=1}^3 \left[ 1 - \sin^2\left(\frac{\pi k_i}{2k_{Ny}}\right) + \frac{2}{15} \sin^4\left(\frac{\pi k_i}{2k_{Ny}}\right) \right]. \quad (3.12)$$

To recover the angle averaged pseudospectrum  $\tilde{P}(k)$  from Eq. (3.10) we use an iterative scheme as described in Jing (2005) with a slight modification: we do not approximate the small scale

<sup>13</sup>For the NGP filter this sum equals 1, and so one recovers the original shot noise term in Eq.(3.1).





**Figure 3.5.:** Power spectrum of the SDSS LRG sample with the bin width  $\Delta k \approx 0.02 h \text{ Mpc}^{-1}$ . The upper solid line shows the best fitting model spectrum and the lower one corresponds to the linearly evolved matter power spectrum of the “concordance” cosmological model multiplied by the square of the bias parameter  $b = 1.95$ . Both of the spectra are convolved with a survey window. The dashed lines represent the “smoothed-out” versions of the above model spectra. The dotted line is the cubic spline fit to the data points.

spectrum by a simple power law, but also allow for the possible running of the spectral index i.e. the parametric shape of the power spectrum is taken to be a parabola in log-log. Since on small scales the power spectrum is dropping fast, the sum over  $\mathbf{n}$  in Eq. (3.10) is converging rather rapidly. In calculations we use only integer vectors with  $|\mathbf{n}| \leq 5$ . The angular average is taken over all the vectors  $\mathbf{k}$  laying in the same  $k$ -space shell with width  $\Delta k$ . The resulting  $\bar{P}$  is taken to be an estimate for the pseudospectrum at the wavenumber  $k_{\text{eff}}$  that corresponds to the average length of the  $k$ -vectors in that shell.

To summarize, our power spectrum calculation consists of the following steps:

1. Determination of the survey selection function i.e. mean underlying number density  $\bar{n}(\mathbf{r})$  (including the survey geometry),
2. Calculation of the overdensity field on a grid using TSC mass assignment scheme,
3. Fourier transformation of the gridded density field,
4. Calculation of the raw 3D power spectrum  $|F(\mathbf{k})|_{\text{FFT}}^2$ ,
5. Subtraction of the shot noise component from the raw spectrum,
6. Recovery of the angle averaged pseudospectrum  $\bar{P}(k)$  using an iterative method of Jing (2005).

We have applied the above described power spectrum calculation method to a multitude of test problems, the results of which can be found in [Hütsi \(2006b\)](#). In Appendix B we show only one example, where we successfully recover the underlying power spectrum of galaxy clusters from the VIRGO Hubble Volume simulations <sup>14</sup>, after applying the selection criteria given in Figs. 3.2 and 3.3.

### 3.4. Power spectrum errors and covariance matrix

We determine power spectrum errors by three different methods:

1. Prescription given by FKP that assumes the underlying density field to be Gaussian. This method also does not treat redshift space distortions. Under those simplifying assumptions the power spectrum variance can be expressed as:

$$\sigma_P^2(k) = \frac{2}{N_k^2} \sum_{\mathbf{k}'} \sum_{\mathbf{k}''} |\tilde{P}(k)Q(\mathbf{k}' - \mathbf{k}'') + S(\mathbf{k}' - \mathbf{k}'')|^2, \quad (3.13)$$

$$Q(\mathbf{k}) = \int d^3r \bar{n}^2(\mathbf{r})w^2(\mathbf{r}) \exp(i\mathbf{k} \cdot \mathbf{r}) \simeq \alpha \sum_j \bar{n}(r_j^s)w^2(r_j^s) \exp(i\mathbf{k} \cdot \mathbf{r}_j^s), \quad (3.14)$$

$$\begin{aligned} S(\mathbf{k}) &= (1 + \alpha) \int d^3r \bar{n}(\mathbf{r})w^2(\mathbf{r}) \exp(i\mathbf{k} \cdot \mathbf{r}) \\ &\simeq \alpha(1 + \alpha) \sum_j w^2(r_j^s) \exp(i\mathbf{k} \cdot \mathbf{r}_j^s). \end{aligned} \quad (3.15)$$

Here the sum is over all the wavevectors  $\mathbf{k}'$  and  $\mathbf{k}''$  populating the same  $k$ -space shell with radius  $k$  and thickness  $\Delta k$ , and  $N_k$  denotes the total number of modes in that shell. Since the direct summation over all the vector pairs  $\mathbf{k}'$  and  $\mathbf{k}''$  is very slow for the wide  $k$ -space shells and  $512^3$  grid we use, a Monte Carlo sum is performed instead. Thus we calculate the average of the quantity  $|\tilde{P}(k)Q(\mathbf{k}' - \mathbf{k}'') + S(\mathbf{k}' - \mathbf{k}'')|^2$  over the random pairs of vectors  $\mathbf{k}'$  and  $\mathbf{k}''$  from the same shell. For the result to converge properly we need on average  $\sim 10^7$  random pairs.

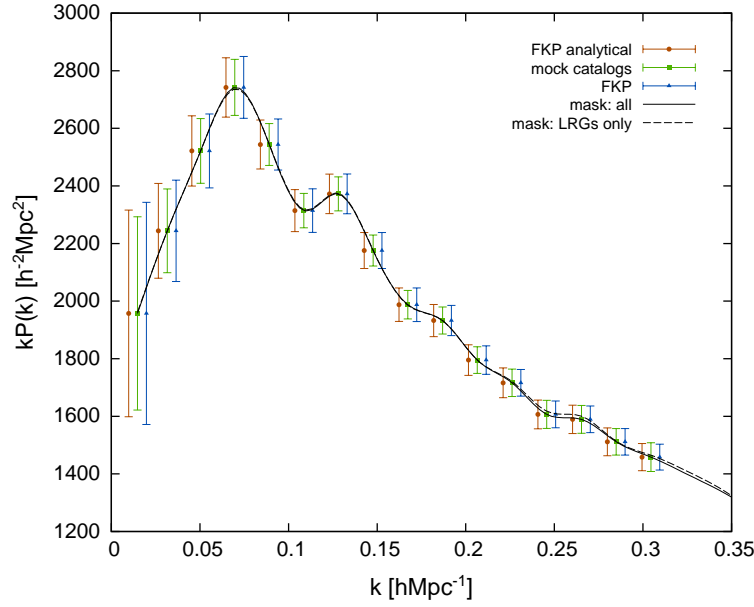
2. The second method is a simple analytical approximation to the first one, also due to FKP (see also [Tegmark et al. 1998](#)). Here the variance is given as:

$$\sigma_P^2(k) = \frac{2\tilde{P}^2(k)}{V_{\text{eff}}V_k}, \quad (3.16)$$

where  $V_k = 4\pi k^2 \Delta k / (2\pi)^3$  is the volume of the  $k$ -space shell and  $V_{\text{eff}}$  is the effective volume given by:

$$V_{\text{eff}} = \frac{\left[ \int d^3r \bar{n}^2(\mathbf{r})w^2(\mathbf{r}) \right]^2}{\int d^3r \bar{n}^4(\mathbf{r})w^4(\mathbf{r}) \left[ 1 + \frac{1}{\bar{n}(\mathbf{r})\tilde{P}(k)} \right]^2}. \quad (3.17)$$

<sup>14</sup><http://www.mpa-garching.mpg.de/Virgo/>



**Figure 3.6.:** The comparison of the different power spectrum error estimates. For clarity slight relative shifts of the data points have been applied. The errorbars resulting from the 1st method are the rightmost ones and the ones from the 3rd method are displayed in the middle. The lines show a cubic spline fits to the data points. The solid line corresponds to the case when all the available galaxy data is used to find the angular mask of the survey, while the dashed line represents the case when LRGs only are used for this purpose.

3. The third method is a Monte Carlo approach that uses 1000 mock catalogs generated in the way described in Appendix C. Here, as we use the 2nd order Lagrangian perturbation theory, we get a good approximation for the mode-mode couplings that are induced during the quasi-nonlinear regime of the evolution of the density fluctuations. Also the large-scale redshift distortions are properly accounted for. In terms of the Halo Model (see Appendix D) we can say that halo-halo clustering term is relatively well approximated. Contributions from the one-halo term can be added later, as these allow an analytic treatment.

The results of the power spectra for the SDSS DR4 LRG sample are shown in Figs. 3.4 and 3.5. In Fig. 3.4 the bin width  $\Delta k \simeq 0.005 h \text{ Mpc}^{-1}$ , while in Fig. 3.5  $\Delta k \simeq 0.02 h \text{ Mpc}^{-1}$ . With different lines we have shown various model spectra, which will be the topic of Sec. 3.6.

The comparison of the power spectrum errorbars calculated in the above described different ways is provided in Fig. 3.6. We see that the various error estimates are in a very good agreement. In the following we will use only the errorbars given by the 3rd method.

So far we have only found the diagonal terms of the covariance matrix. In order to answer the question of how strongly different power spectrum bins are correlated, we have to go a step further, and try to estimate the full covariance matrix.

The FKP result for the full covariance matrix,  $C_{ij}$ , is a simple generalization of the Eq. (3.13):

$$C_{ij} = \frac{2}{N_{k'} N_{k''}} \sum_{\mathbf{k}'} \sum_{\mathbf{k}''} \left| \tilde{P}\left(\frac{k_i + k_j}{2}\right) Q(\mathbf{k}' - \mathbf{k}'') + S(\mathbf{k}' - \mathbf{k}'') \right|^2, \quad (3.18)$$

where the  $k$ -vectors  $\mathbf{k}'$  and  $\mathbf{k}''$  lie in shells with width  $\Delta k$  and radii  $k_i$  and  $k_j$ , respectively. The FKP approach, as mentioned above, does not treat mode couplings arising from the nonlinear evolution and also from the redshift space distortions. Linear redshift space distortions can be, in principle, included into the FKP estimate for the covariance matrix. One can generalize the results presented in the Appendix of [Zaroubi & Hoffman \(1996\)](#), where the covariance matrix for the Fourier modes has been found. Since linear redshift distortions applied on a Gaussian field do not change the Gaussianity property, one can still use the result from the Appendix B of FKP that relates the power spectrum covariance matrix to the covariance matrix of the Fourier modes. Also one has to add the shot noise terms to the result of [Zaroubi & Hoffman \(1996\)](#). We have carried out this exercise, leading us to the high dimensional integrals (up to 12 dim.) that turn out to be too time consuming to solve in practice. As from the mock catalogs we can hopefully obtain more realistic estimate for the covariance matrix <sup>15</sup> we have not followed this path any further.

The results for the covariance matrix calculation are given in Fig. 3.7. Here the left hand column shows the covariance and the right hand column respective correlation matrices:

$$r_{ij} = \frac{C_{ij}}{\sqrt{C_{ii}C_{jj}}}. \quad (3.19)$$

The power spectrum binning is the same as shown in Fig. 3.5 i.e.  $\Delta k \simeq 0.02 h \text{ Mpc}^{-1}$ . The top row represents the results from Eq. (3.18), while the middle row the ones from mock catalogs. Although the diagonal terms of the covariance matrices in the 1st and 2nd row are in a very good agreement (see Fig. 3.6), the off-diagonal components differ strongly. This can be explained as the result of the extra mode-mode couplings that are not accounted for by the FKP approach. We see that even well separated power spectrum bins can be correlated at 30...40% level. The bottommost row in Fig. 3.7 represents the nonlinear contribution to the covariance matrix arising from the 1-halo term (see Appendix D). We see that this contribution is subdominant at the scales of interest to us <sup>16</sup>.

In the following calculations we mostly use the covariance matrix given in the middle row of Fig. 3.7.<sup>17</sup>

### 3.5. Relation to the true spectrum

Since masking in real space is equivalent to convolution in Fourier space, our measured power spectrum  $\tilde{P}$  is actually a convolution of the real spectrum  $P$  with a survey window (see e.g. FKP):

$$\tilde{P}(\mathbf{k}) = \int \frac{d^3k'}{(2\pi)^3} P(\mathbf{k}) |W(\mathbf{k} - \mathbf{k}')|^2, \quad (3.20)$$

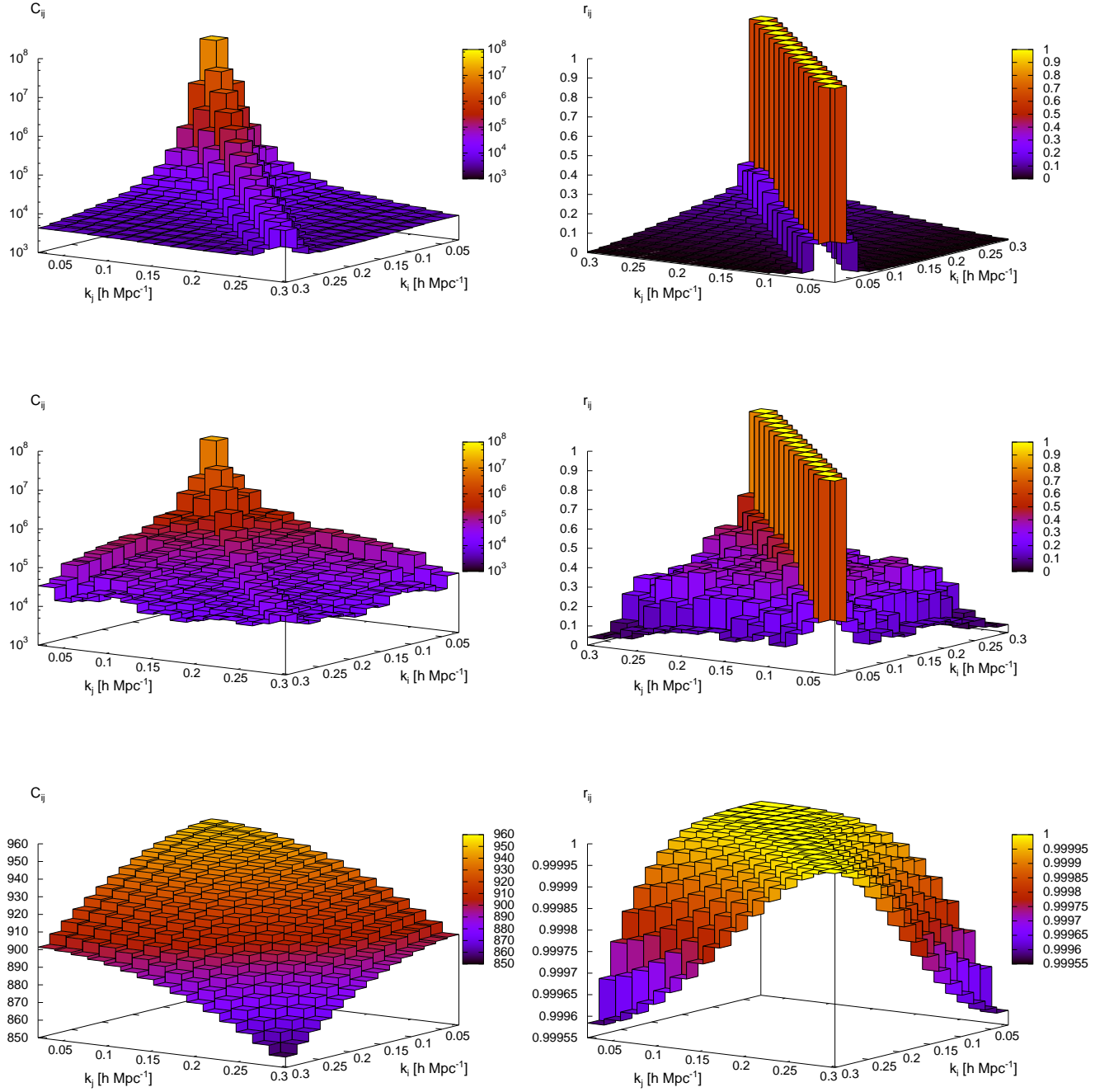
where

$$W(\mathbf{k}) = \int d^3r \tilde{n}(\mathbf{r}) w(\mathbf{r}) \exp(i\mathbf{k} \cdot \mathbf{r}), \quad (3.21)$$

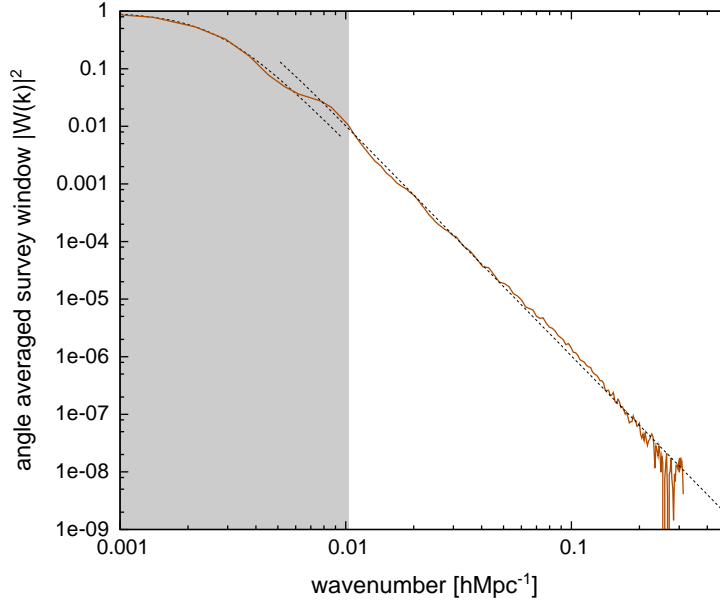
<sup>15</sup>Since now we are also able to handle quasi-nonlinear mode-mode couplings.

<sup>16</sup>In calculating this contribution to the covariance matrix we have taken the best fit model parameters as obtained in Sec. 3.6. The smallness of this term is caused by the high value of the parameter  $M_0$  i.e. majority of the “occupied” halos contain only one LRG.

<sup>17</sup>This matrix along with the power spectrum results in Fig. 3.5 is also given in a tabular form in Appendix H.



**Figure 3.7.:** Covariance (left column) and correlation matrices (right column). Top row represents the results from FKP prescription (see Eq. (3.18)) and the middle row the ones from 1000 mock catalogs. The last row displays the nonlinear contribution due to the 1-halo term.



**Figure 3.8.:** Isotropized survey window. Here the normalization is taken such that  $|W(0)| = 1$ . Light gray stripe marks the region where the window is above 1% of its maximum value of 1. Dashed lines show approximations discussed in the text.

and the survey window  $|W(\mathbf{k})|^2$  is normalized as follows:

$$\int \frac{d^3k}{(2\pi)^3} |W(\mathbf{k})|^2 = 1. \quad (3.22)$$

The angle averaged survey window  $|W(k)|^2$  is plotted in Fig. 3.8. Here the core part of the window is well approximated by the functional form:

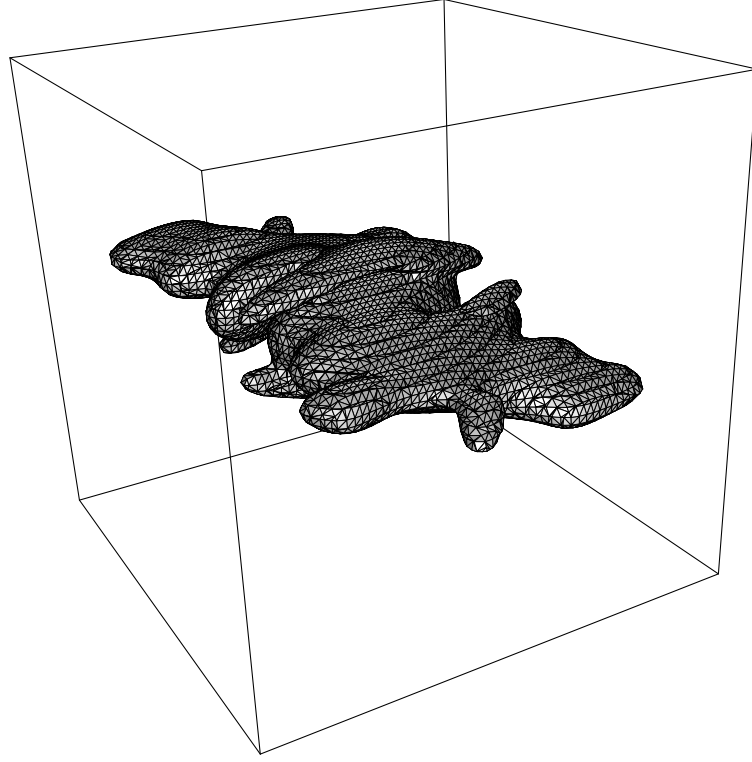
$$|W(k)|^2 = \frac{1}{1 + \left(\frac{k}{a}\right)^2 + \left(\frac{k}{b}\right)^4}, \quad (a \simeq 0.0030, b \simeq 0.0028), \quad (3.23)$$

and asymptotic wings are close to the power law with spectral index  $-4$ . These approximations are shown with dashed lines in Fig. 3.8. With the gray shaded stripe we have marked the scales where  $|W(k)|^2$  is above 1% of its maximum value. This stripe just serves as a rough guide to the effective width of the survey window and it is also shown in many of the following figures.

Since the survey geometry of the analyzed SDSS LRG sample is far from being spherically symmetric, an isotropized window in Fig. 3.8 gives only a poor representation of the true 3D window, which is displayed as an isosurface corresponding to the isovalue of 0.01 in Fig. 3.9.

In order to compare theoretical models to the measured power spectrum we have to take into account the smearing effects caused by the survey window. Using Eq. (3.20) we can express an isotropized power spectrum as:

$$\tilde{P}(k) = \int \frac{d\Omega_{\mathbf{k}}}{4\pi} \tilde{P}(\mathbf{k}) = \int dk' k'^2 P(k') K(k', k), \quad (3.24)$$



**Figure 3.9.:** 3D survey window embedded in a box with a side length of  $0.04 h \text{ Mpc}^{-1}$ . Here the isosurface corresponding to 1% of the maximum value of the window is shown. Note the symmetry of the window,  $|W(\mathbf{k})|^2 = |W(-\mathbf{k})|^2$ , as expected when taking a modulus of the Fourier transform of a real 3D scalar function.

where the coupling kernels <sup>18</sup>:

$$K(k', k) = K(k, k') = \frac{1}{2\pi^2} \int \frac{d\Omega_{\mathbf{k}}}{4\pi} \int \frac{d\Omega_{\mathbf{k}'}}{4\pi} |W(\mathbf{k} - \mathbf{k}')|^2. \quad (3.25)$$

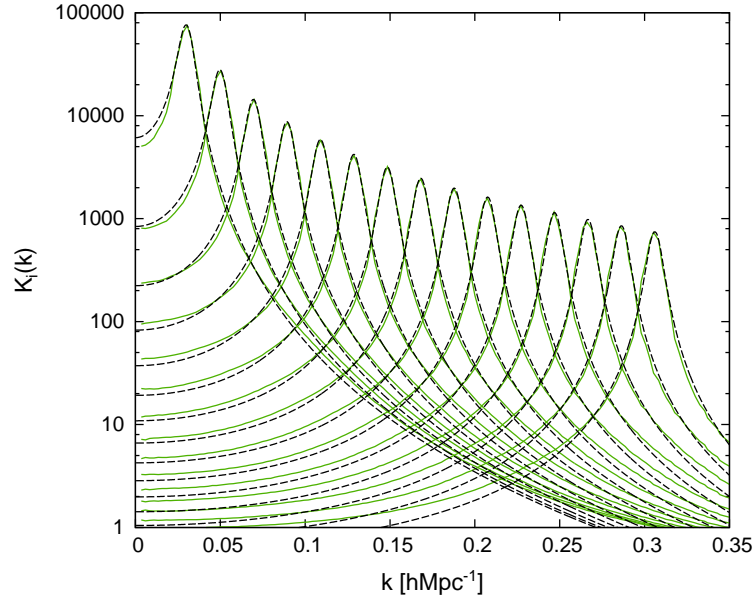
Numerically evaluated coupling kernels along with the analytical approximations (see Appendix E) for the analyzed galaxy sample are presented in Fig. 3.10. Here the solid lines correspond to the numerical results and the dashed ones represent an analytical approximation. We have used the notation  $K_i(k) \equiv K(k_i, k)$  where  $k_i$  denote the central values of the power spectrum bins shown in Fig. 3.5.

### 3.6. Model spectra

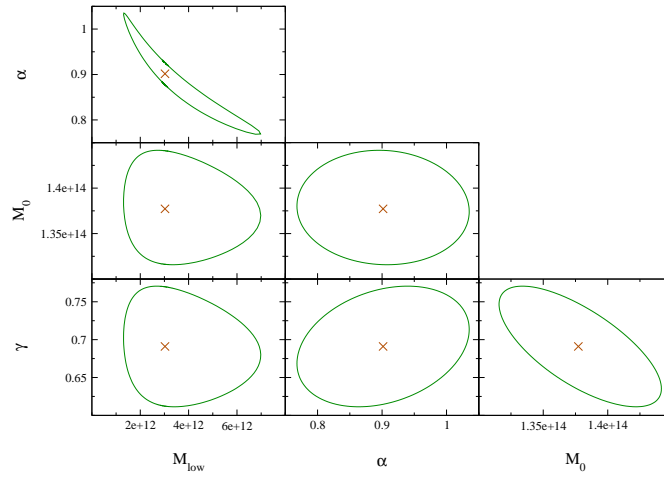
It is well known that redshift space distortions and nonlinear effects modify simple linear spectra. In order to treat these effects we make use of the very successful analytical model – the Halo Model. For a nice review we refer the reader to Cooray & Sheth (2002) (see also Seljak 2000).

<sup>18</sup>We prefer to use “coupling kernels” instead of the more common “window functions” since the word “window” has already been used to mean the modulus square of the Fourier transform of the weighted survey volume.





**Figure 3.10.:** Coupling kernels  $K_i(k) \equiv K(k_i, k)$  for the power spectrum bins  $k_i$  shown in Fig. 3.5. Numerically evaluated kernels are shown with solid lines. The dashed lines correspond to the fitting functions given in Appendix E.



**Figure 3.11.:**  $1\sigma$  error contours for the free model parameters. Best fit parameter values are marked with crosses.

The details of the model we use are presented in Appendix D. The model introduces four free parameters:  $M_{\text{low}}$ ,  $\alpha$ ,  $M_0$  and  $\gamma$ . Here  $M_{\text{low}}$  is the lower cutoff of the halo mass i.e. below that mass halos are assumed to be “dark”.  $\alpha$  and  $M_0$  are the parameters of the mean of the halo occupation distribution  $\langle N|M \rangle$ , which gives the average number of galaxies per halo with mass  $M$ . We take  $\langle N|M \rangle$  to be a simple power law:

$$\langle N|M \rangle = \left( \frac{M}{M_0} \right)^\alpha. \quad (3.26)$$

The last parameter,  $\gamma$ , is the amplitude factor for the virial velocities of galaxies inside dark matter halos. One dimensional velocity dispersion of the galaxies inside a halo with mass  $M$  is taken to follow the scaling of the isothermal sphere model:

$$\sigma = \gamma \sqrt{\frac{GM}{2R_{\text{vir}}}}, \quad (3.27)$$

where  $R_{\text{vir}}$  is the virial radius of the halo.

For the model fitting we have used Levenberg-Marquardt method as described in [Press et al. \(1992\)](#) with modifications (described in Appendix F) that allow us to incorporate correlations between the data points. As the input data we take the power spectrum estimates given in Fig. 3.5. The covariance matrix used is the one shown in the middle row of Fig. 3.7. We also perform fits where we use one additional power spectrum bin on a larger scale (not shown in Fig. 3.5). All of this data is given in a tabular form in Appendix H. The transfer functions needed for the linear spectra are taken from [Eisenstein & Hu \(1998\)](#). There the authors also provide transfer function fits where the baryonic acoustic oscillations have been removed. We use these “smoothed out” transfer functions in order to assess the significance of the oscillatory features we see in the data. Throughout this paper we have kept cosmology fixed to the best fit WMAP “concordance” model ([Spergel et al. 2003](#)). The implications for the cosmology, and especially for the dark energy equation of state parameter, are planned to be worked out in the future paper.

As the cosmology is kept fixed, we have only four free parameters. In order to eliminate some of the degeneracies between the parameters we have imposed one additional constraint. Namely, we have demanded that the resulting number of galaxies should agree with the one that is observed with the relative error of 1% i.e.  $(51,763 \pm 518)^{19}$ . The resulting  $1\sigma$  error “ellipses” for the free parameters are shown in Fig. 3.11. The “ellipses” appear deformed since instead of  $M_{\text{low}}$  and  $M_0$  we have fitted  $\log(M_{\text{low}})$  and  $\log(M_0)$ . With crosses we have marked the best fit values:  $M_{\text{low}} \simeq 3 \cdot 10^{12} h^{-1} M_\odot$ ,  $\alpha \simeq 0.9$ ,  $M_0 \simeq 1.4 \cdot 10^{14} h^{-1} M_\odot$  and  $\gamma \simeq 0.7$ . The model spectra corresponding to these best fit parameters are shown in Figs. 3.4 and 3.5. In both figures we have also given the simple linear spectra multiplied by the square of the bias parameter  $b = 1.95$ . In Fig. 3.4 we have additionally demonstrated the effect of the window convolution. There the dashed lines correspond to the unconvolved case. In Fig. 3.5 along with the “wiggly” spectra we have shown their “smoothed-out” counterparts. Using all the 16 power spectrum bins (the 1st not shown in Fig. 3.5) plus an additional constraint on the total number of galaxies, resulting in  $17 - 4 = 13$  independent degrees of freedom, we obtain  $\chi^2$  values of 8.8 and 19.9 for the

---

<sup>19</sup>The  $1\sigma$  Poisson error in this case would be 228. The large-scale structure amplifies the variability in the number of objects and a factor of a few increase above the Poissonian case seems to be reasonable.

“wiggly” and “smoothed”<sup>20</sup> models, respectively. So the models with oscillations are favored by  $3.3\sigma$  over their “smoothed-out” counterparts.<sup>21</sup> Since both models have the same number of free parameters, and if additionally the assumption of Gaussianity is valid, the Bayesian approach should also give similar results. Actually, Bayesian results should favor “wiggly” models even more, since prior weight for these should probably be taken higher (assuming the knowledge of the other experimental results).

### 3.7. Determination of the acoustic scale

To measure the scale of the acoustic oscillations we divide the spectrum shown in Fig. 3.5 with the best fitting “smoothed” spectrum. The result of this procedure is given in the upper panel of Fig. 3.12. There the solid line shows a cubic spline fit to the data points and the long-dashed line corresponds to the best fitting model spectrum also shown in Fig. 3.5. The above data is fitted with a parametric form:

$$f(x) = 1 + c_1 \cdot \sin(c_2 \cdot x) \exp \left[ \left( -\frac{x}{c_3} \right)^{c_4} \right]. \quad (3.28)$$

Again we use the Levenberg-Marquardt method with the data covariance matrix obtained from mock catalogs. After marginalizing over the other parameters we find the best fitting value of  $(105.4 \pm 2.3) h^{-1} \text{ Mpc}$  for the parameter  $c_2$ .<sup>22</sup> The best fitting member of the parametric family in Eq. (3.28) is shown with short-dashed lines in the upper panel of Fig. 3.12. Using FKP covariance matrix instead gives an acoustic scale of  $(105.4 \pm 2.8) h^{-1} \text{ Mpc}$ .

The sinusoidal modulation in the power spectrum is a pure consequence of the adiabatic initial conditions. By relaxing this assumption and fitting with a more general functional form:

$$f(x) = 1 + c_1 \cdot \sin(c_2 \cdot x + c_3) \exp \left[ \left( -\frac{x}{c_4} \right)^{c_5} \right]. \quad (3.29)$$

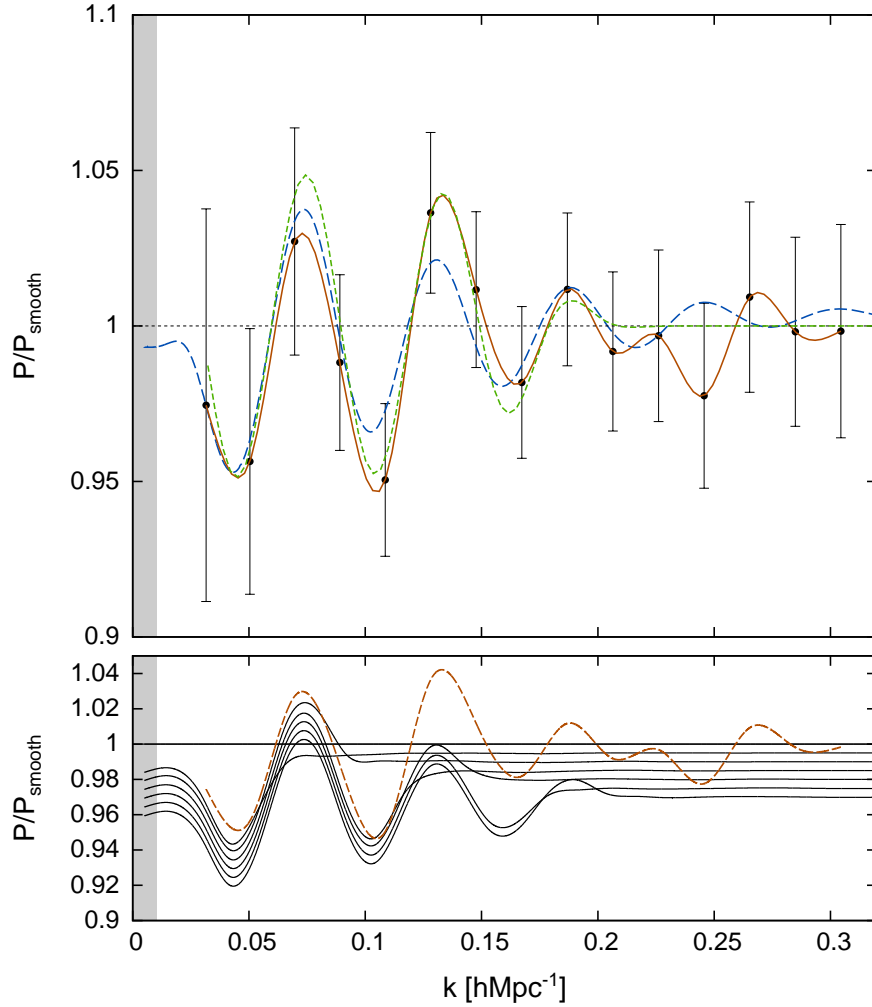
instead, we get the following value for the acoustic scale:  $(103.0 \pm 7.6) h^{-1} \text{ Mpc}$ . In case of the FKP covariance matrix the corresponding value is  $(103.1 \pm 9.1) h^{-1} \text{ Mpc}$ .

Eisenstein et al. (2005) determine various distance scales (like  $D_v$ , which is a certain mixture of the comoving distances along and perpendicular to the line of sight (see their Eq.(2))) and their ratios, using SDSS LRGs in combination with the constraints from other cosmological sources. The typical relative accuracy of these measurements is  $\sim 4\%$ , which might seem to be significantly poorer than the accuracy of the acoustic scale measurement,  $(105.4 \pm 2.3) h^{-1} \text{ Mpc}$  i.e.  $\sim 2\%$ , presented in this paper. This apparent inconsistency can be attributed to the fact that in

<sup>20</sup>The best fit  $M_{\text{low}}$ ,  $\alpha$ ,  $M_0$  and  $\gamma$  for the “smoothed” models differ slightly from the values quoted above for the “wiggly” spectra.

<sup>21</sup>Dropping the first power spectrum bin the obtained  $\chi^2$  values are 5.0 and 16.5. 5.0 is an anomalously low value of  $\chi^2$  for 12 degrees of freedom. (One would expect  $\chi^2 \simeq 12 \pm 5$ .) In fact, if we would have used the simple FKP covariance matrix instead of the one obtained from the mock catalogs, the resulting  $\chi^2$  values would be even lower: 2.9 and 8.5, respectively. This might hint that the 2nd order Lagrangian approach, although very successful, might still have problems of capturing some extra mode-mode couplings.

<sup>22</sup>Here and in the following all the errors refer to the  $1-\sigma$  level. Values for the other parameters are as follows:  $c_1 = (4.9 \pm 2.1) \cdot 10^{-2}$ ,  $c_3 = (0.176 \pm 0.023) h \text{ Mpc}^{-1}$ ,  $c_4 = (7 \pm 17)$ .



**Figure 3.12.:** Upper panel: Power spectrum from Fig. 3.5 divided by the best fitting “smoothed” spectrum. Solid line shows a cubic spline fit to the data points and long-dashed line corresponds to the best “wiggly” model. The short-dashed line represents the most favorable fit from the parametric family of Eq. (3.28). Lower panel: Various input power spectra used to calculate the two-point correlation function. The dashed line is the cubic spline fit from the upper panel. The solid lines represent a transition sequence from the best fitting “wiggly” model to the best “smoothed” model. In each step we have erased more and more oscillatory features. For clarity slight vertical shifts have been introduced.

our analysis, as stated above, we have kept the cosmology fixed to the WMAP “concordance” model, whereas [Eisenstein et al. \(2005\)](#) estimates include the extra uncertainties due to the imperfect knowledge of the various cosmological parameters. Of course, the given length of the acoustic scale,  $(105.4 \pm 2.3) h^{-1} \text{ Mpc}$ , can be easily transformed in order to accommodate other preferences for the background cosmology. We also note that the use of the parametric form in Eq. (3.28) might be too restrictive, since the acoustic modulation in the case of adiabatic models can be only approximately described as a damped sinusoidal wave ([Eisenstein & Hu 1998](#)). For this reason the given sound horizon constraint should not be used in cosmological parameter studies. Instead one should directly use the measured power spectrum in combination with the parametrized models that are physically well motivated.

### 3.8. Correlation function analysis

We determine the two-point correlation function of the SDSS LRGs using the edge-corrected estimator given by [Landy & Szalay \(1993\)](#):

$$\xi(r) = \frac{DD - 2DR + RR}{RR}, \quad (3.30)$$

which has minimal variance for a Poisson process. Here DD, DR and RR represent the respective normalized data-data, data-random and random-random pair counts in a given distance range. Random catalogs were generated with 25 times the number of objects in the main catalog. We calculated correlation function for 10 bins ( $r_i$ ,  $i = 1 \dots 10$ ) in the pair distance range of  $60 \dots 160 h^{-1} \text{ Mpc}$ . The errors were estimated by a “jackknife” technique. For this purpose we divided the full sample into 22 separate regions each containing  $\sim 2350$  galaxies (see Fig. 3.3). The two-point function was calculated 22 times, each time omitting one of the regions. Denoting the resulting estimates as  $\xi_j(r_i)$ , ( $j = 1 \dots 22$ ), the “jackknife” estimate for the variance reads as (see e.g. [Lupton 1993](#)):

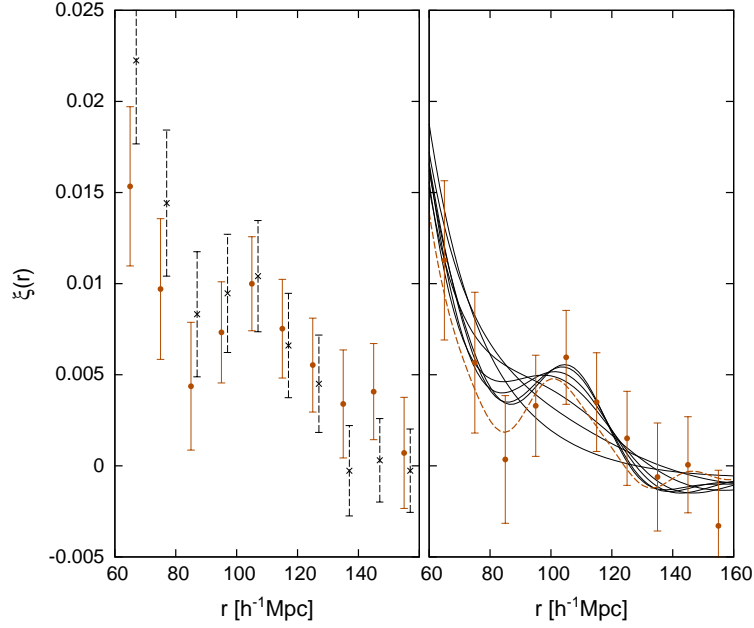
$$\sigma_{\xi}^2(r_i) = \frac{N-1}{N} \sum_{j=1}^N [\xi_j(r_i) - \bar{\xi}(r_i)]^2, \quad (3.31)$$

$$\bar{\xi}(r_i) = \frac{1}{N} \sum_{j=1}^N \xi_j(r_i), \quad (3.32)$$

where in our case  $N = 22$ . The results of this calculation are presented in the left panel of Fig. 3.13. With the crosses and dashed-line errorbars we have also shown the two-point function as determined by [Eisenstein et al. \(2005\)](#). We see that in general our results agree reasonably well with their calculations.

It would be interesting to study how the oscillations in the observed power spectrum transform into the peak in the two-point correlation function seen at the scale of  $\sim 110 h^{-1} \text{ Mpc}$ . For this purpose we use the cubic spline fit shown in Fig. 3.12 and extend it outside of the observed range by smoothly joining it to the power spectrum of the best fitting “smoothed-out” model. The correlation function is now simply calculated as the Fourier transform of the power spectrum.<sup>23</sup>

<sup>23</sup>To be precise, in redshift space the two-point correlation function and power spectrum are not anymore exact



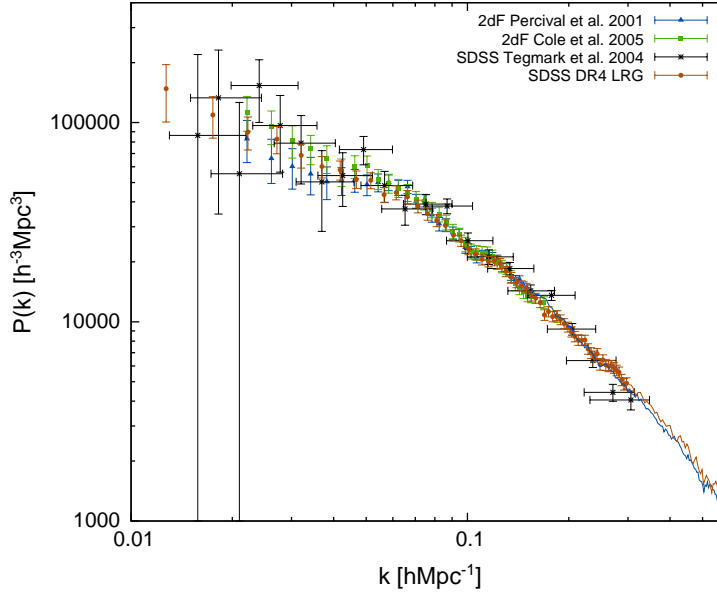
**Figure 3.13.:** Left panel: Two-point correlation functions as determined in this paper (circles with solid lines) and by Eisenstein et al. (2005). Right panel: Correlation functions corresponding to the models shown in the lower panel of Fig. 3.12 in comparison to the one obtained directly from the data. Here all the data points have been lowered by 0.0035.

The resulting correlation function is plotted with a dashed line in the right panel of Fig. 3.13. To study the significance of the oscillatory features in the power spectrum in relation to the observed peak in the correlation function, we have calculated correlation functions for several models that have oscillations “switched off” at various scales. The spectra of these models are shown with solid lines in the lower panel of Fig. 3.12, where for the sake of clarity we have introduced slight vertical shifts between the curves, so that the scales where the transition to the featureless spectrum takes place, are easily visible. The corresponding correlation functions are given with solid lines in the right hand panel of Fig. 3.13. As expected, we see how the peak in the correlation function is getting broader and also decreasing in amplitude as we erase more and more features in the power spectrum. This clearly demonstrates the importance of many of the up-downs in the power spectrum to produce a relatively sharp feature in the two-point correlation function.

In order to achieve good agreement we have lowered all the data points by 0.0035 in the right hand panel of Fig. 3.13. Similar shifts were also suggested in Eisenstein et al. (2005) in order to get better match to the theoretical models. A 0.0035 shift in  $\xi$  translates to the 0.175% shift in the mean density. Thus, if one wishes to determine the amplitude of the correlation function correctly at those large scales, one has to determine the survey selection function with a very high precision, which in practice is very difficult to achieve. By using model spectra that have more

---

Fourier transforms of each other. Nevertheless, we think that this simplified exercise is still useful. Also, as the correlation function estimator in Eq. (3.30) is an edge-corrected estimator, we use an unconvolved model spectra here.



**Figure 3.14.:** The comparison of spectra from different surveys.

large scale power than the “concordance” cosmology predicts (as might be suggestive from Fig. 3.4), we are in fact able to match the amplitude of the correlation function without any additional vertical shifts. Here we try to avoid making any definite conclusions. The behavior of the power spectrum on the largest scales is an extremely interesting topic on its own and there exist much better methods than the direct Fourier approach to investigate these issues (see e.g. Tegmark et al. 1998).

### 3.9. Comparison with the other surveys

In this section we compare our power spectrum measurements with the ones obtained by Percival et al. (2001) and Cole et al. (2005) for the 2dF redshift survey and by Tegmark et al. (2004a) for the SDSS main galaxy sample. The results of this comparison are provided in Figs. 3.14 and 3.15. Since Fig. 3.14 is extremely busy, we have also given a variant of it where we have omitted the errorbars. The amplitudes of the SDSS main and 2dF spectra have been freely adjusted to match the clustering strength of the SDSS LRGs. The corresponding bias parameters with respect to the SDSS LRGs are 0.53, 0.61 and 0.50 for the 2dF sample analyzed by Percival et al. (2001), for the one analyzed by Cole et al. (2005), and for the SDSS main sample, respectively. Percival et al. (2001) also provide power spectrum measurements for  $k \gtrsim 0.15 h \text{ Mpc}^{-1}$  but without errorbars. These small-scale measurements along with our SDSS LRG results are shown with solid lines in Fig. 3.14.

In general the shapes of the spectra agree remarkably well. Of course one has to keep in mind that here, with the only exception of Tegmark et al. (2004a) results, the power spectrum bins are highly correlated. Also Tegmark et al. (2004a) measurements are corrected for the redshift space distortions.



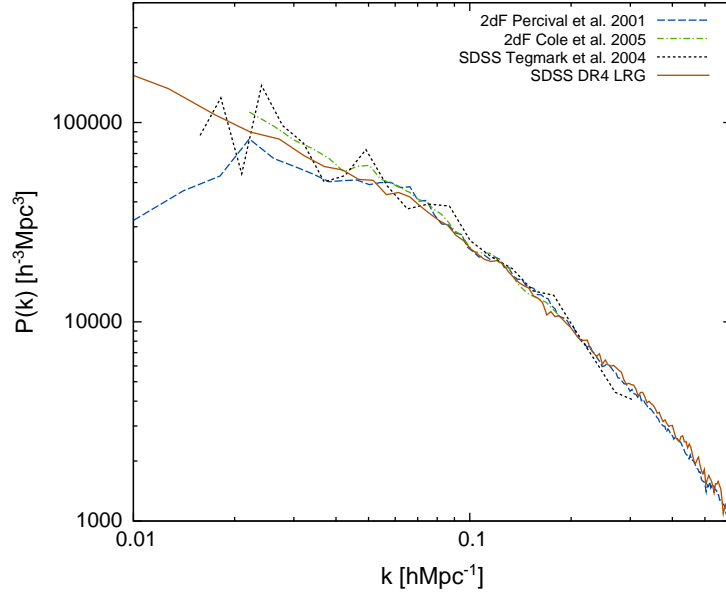


Figure 3.15.: The same as Fig. 3.14 with the errorbars omitted.

### 3.10. Discussion and Conclusions

In this paper we have calculated the redshift-space power spectrum of the SDSS DR4 LRG sample, finding evidence for a series of acoustic features down to the scales of  $\sim 0.2 h \text{ Mpc}^{-1}$ . It turns out that models with the baryonic oscillations are favored by  $3.3\sigma$  over their “smoothed-out” counterparts without any oscillatory behavior. Using the obtained power spectrum we predict the shape of the spatial two-point correlation function, which agrees very well with the one obtained directly from the data. Also, the directly calculated correlation function is consistent with the results obtained by Eisenstein et al. (2005). We have made no attempts to put constraints on the cosmological parameters, rather we have assumed in our analysis the “concordance” cosmological model. The derived acoustic scale  $(105.4 \pm 2.3) h^{-1} \text{ Mpc}$  agrees well with the best-fit WMAP “concordance” model prediction of  $\approx 106.5 h^{-1} \text{ Mpc}$ .

The existence of the baryonic features in the galaxy power spectrum is very important, allowing one (in principle) to obtain Hubble parameter  $H$  and angular diameter distance  $d_A$  as a function of redshift, this way opening up a possibility to constrain properties of the dark energy (Hu & Haiman 2003). The currently existing biggest redshift surveys, which are still quite shallow, do not yet provide enough information to carry out this project fully. On the other hand, it is extremely encouraging that even with the current generation of redshift surveys we are already able to see the traces of acoustic oscillations in the galaxy power spectrum, showing the great promise for the dedicated future surveys like K.A.O.S. We have seen that acoustic features seem to survive at mildly nonlinear scales ( $k \gtrsim 0.1 h \text{ Mpc}^{-1}$ ), which is in agreement with the results of the recent N-body simulations (Springel et al. 2005, Seo & Eisenstein 2005). In order to fully exploit available information one needs a complete understanding of how nonlinear effects influence these features. Nonlinear bias and redshift space distortions also add extra complications. In general redshift-space distortions, biasing and nonlinear evolution do not create any oscilla-

tory modulation in the power spectrum and so acoustic features should be readily observable. So far there have been only a few works studying these important issues (e.g. [Springel et al. \(2005\)](#), [Seo & Eisenstein \(2005\)](#), [White \(2005\)](#)) and probably it is fair to say that currently we really do not have a full theoretical description of them. In our paper we have modeled the above mentioned effects using the results from the 2nd order Lagrangian perturbation theory in combination with the Halo Model. Although these models are very successful in capturing many important aspects of the structure formation, one has to keep in mind that they are still approximations.

The bare existence of the baryonic oscillations in the galaxy power spectrum tells us something important about the underlying cosmological model and the mechanism of the structure formation. First, it confirms the generic picture of the gravitational instability theory where the structure in the Universe is believed to be formed by the gravitational amplification of the small perturbations laid down in the early Universe. Under the linear gravitational evolution all the density fluctuation modes evolve independently i.e. all the features in the power spectrum will be preserved. And certainly, we are able to identify features in the low redshift galaxy power spectrum that correspond to the fluctuations seen in the CMB angular power spectrum (which probes redshifts  $z \sim 1100$ ), providing strong support for the above described standard picture of the structure formation. Actually, we can also probe scales that are inaccessible for the CMB studies due to the strong damping effects and steeply rising influence of the secondary anisotropies, reaching effectively the wavenumbers that correspond to the 6th-7th peak in the CMB angular power spectrum. Second, the ability to observe baryonic features in the low redshift galaxy power spectrum demands rather high baryonic to total matter density ratio. In [Blanchard et al. \(2003\)](#) it has been shown that it is possible to fit a large body of observational data with an Einstein–de Sitter type model if one adopts low value for the Hubble parameter and relaxes the usual assumptions about the single power law initial spectrum. In the light of the results obtained in our paper these models are certainly disfavored due to the fact that the high dark matter density completely damps the baryonic features. And finally, purely baryonic models are also ruled out since for them the expected acoustic scale would be roughly two times larger than observed here <sup>24</sup>. So the data seems to demand a weakly interacting nonrelativistic matter component and all the models that try to replace this dark matter component with something else e.g. modifying the laws of gravity might have severe difficulties to fit these new observational constraints.

---

<sup>24</sup>For a clear discussion of this see Daniel Eisenstein's home page  
<http://cmb.as.arizona.edu/~eisenste/acousticpeak/>



# 4. Power spectrum of the SDSS luminous red galaxies: constraints on cosmological parameters

G. Hütsi

*Astronomy & Astrophysics*, submitted, [Hütsi \(2006\)](#)

## Abstract

In this paper we determine the constraints on cosmological parameters using the CMB data from the WMAP experiment together with the recent power spectrum measurement of the SDSS Luminous Red Galaxies (LRGs). Specifically, we focus on spatially flat, low matter density models with adiabatic Gaussian initial conditions. The spatial flatness is achieved with an additional quintessence component whose equation of state parameter  $w_{\text{eff}}$  is taken to be independent of redshift. We do not allow any massive neutrino contribution and also the influence of the gravitational waves on the CMB is taken to be negligible. The analysis is carried out separately for two cases: (i) using the acoustic scale measurements as presented in [Hütsi \(2006a\)](#), (ii) using the full SDSS LRG power spectrum and its covariance matrix. We are able to obtain a very tight constraint on the Hubble parameter:  $H_0 = 70.8^{+1.9}_{-1.8}$  km/s/Mpc, which helps in breaking several degeneracies between the parameters and allows us to determine the low redshift expansion law with much higher accuracy than available from the WMAP + HST data alone. The positive deceleration parameter  $q_0$  is found to be ruled out at  $5.5\sigma$  confidence level.

## 4.1. Introduction

Since the flight of the COBE <sup>1</sup> satellite in the beginning of 90's the field of observational cosmology has witnessed an extremely rapid development. The data from various Cosmic Microwave Background (CMB) experiments (WMAP <sup>2</sup> ([Bennett et al. 2003](#)), COBE ([Smoot et al. 1992](#)), ARCHEOPS <sup>3</sup> ([Benoît et al. 2003](#)), BOOMERANG <sup>4</sup> ([Netterfield et al. 2002](#)), MAXIMA <sup>5</sup> ([Hanany et al. 2000](#)), CBI <sup>6</sup> ([Pearson et al. 2003](#)), VSA <sup>7</sup> ([Scott et al. 2003](#)), DASI <sup>8</sup> ([Halverson et al. 2002](#)))

---

<sup>1</sup><http://lambda.gsfc.nasa.gov/product/cobe/>

<sup>2</sup><http://map.gsfc.nasa.gov/>

<sup>3</sup><http://www.archeops.org/>

<sup>4</sup><http://cmb.phys.cwru.edu/boomerang/>

<sup>5</sup><http://cfpa.berkeley.edu/group/cmb/>

<sup>6</sup><http://www.astro.caltech.edu/~tjp/CBI/>

<sup>7</sup><http://www.mrao.cam.ac.uk/telescopes/vsa/>

<sup>8</sup><http://astro.uchicago.edu/dasi/>

etc.), supernova surveys (SCP <sup>9</sup> (Perlmutter et al. 1999), High-Z SN Search <sup>10</sup> (Riess et al. 1998)) and large galaxy redshift surveys (SDSS <sup>11</sup> (York et al. 2000), 2dFGRS <sup>12</sup> (Colless et al. 2001)) has lead us to the cosmological model that is able to accommodate almost all the available high quality data– the so-called “concordance” model (Bahcall et al. 1999, Spergel et al. 2003). Useful cosmological information has also been obtained from the Ly- $\alpha$  forest, weak lensing, galaxy cluster, and large-scale peculiar velocity studies. It is remarkable that this diversity of observational data can be fully explained by a cosmological model that in its simplest form has only 5 – 6 free parameters (Liddle 2004, Tegmark et al. 2004b). As the future data sets will be orders of magnitude larger, leading to the extremely small statistical errors, any further progress is possible only in case we fully understand various systematic uncertainties that could potentially bias our conclusions about the underlying cosmology. As such, one should try to use observables that are least sensitive to the theoretical uncertainties, contaminating foregrounds etc. Currently the “cleanest” constraints on cosmological models are provided by the measurements of the angular power spectrum of the CMB. Since the underlying linear physics is well understood (see e.g. Hu 1995, Dodelson 2003) we have a good knowledge of how the angular position and amplitude ratios of the acoustic peaks depend on various cosmological parameters. However, the CMB data alone is able to provide accurate measurements of only a few combinations of the cosmological parameters. In order to break the degeneracies between the parameters one has to complement the CMB data with additional information from other independent sources e.g. the data from the type Ia supernovae, large-scale structure, or the Hubble parameter measurements. In fact, the well understood physical processes responsible for the prominent peak structure in the CMB angular power spectrum are also predicted to leave imprints on the large-scale matter distribution. Recently the analysis of the spatial two-point correlation function of the Sloan Digital Sky Survey (SDSS) Luminous Red Galaxy (LRG) sample (Eisenstein et al. 2005), and power spectra of the 2dF (Cole et al. 2005) and SDSS LRG (Hütsi 2006a) redshift samples, have lead to the detection of these acoustic features, providing the clearest support for the gravitational instability picture, where the large-scale structure of the Universe is believed to arise through the gravitational amplification of the density fluctuations laid down in the very early Universe.

In the current paper we work out the constraints on cosmological parameters using the SDSS LRG power spectrum as determined by Hütsi (2006a). In order to break the degeneracies between the parameters we complement our analysis with the data from other cosmological sources: the CMB data from the WMAP, and the measurement of the Hubble parameter by the HST Key Project <sup>13</sup>. We focus our attention on simple models with Gaussian adiabatic initial conditions. We further assume spatial flatness, and also negligible massive neutrino and gravitational wave contributions. This leads us to the models with 6 free parameters: total matter and baryonic matter density parameters:  $\Omega_m$  and  $\Omega_b$ , the Hubble parameter  $h$ , the optical depth to the last-scattering surface  $\tau$ , the amplitude  $A_s$  and spectral index  $n_s$  of the scalar perturbation spectrum. <sup>14</sup> This minimal set is extended with the constant dark energy effective equation of state parameter  $w_{\text{eff}}$ . We carry out our analysis in two parts. In the first part we use only the mea-

---

<sup>9</sup><http://supernova.lbl.gov/>

<sup>10</sup><http://cfa-www.harvard.edu/oir/Research/supernova/HighZ.html>

<sup>11</sup><http://www.sdss.org/>

<sup>12</sup><http://www.mso.anu.edu.au/2dFGRS/>

<sup>13</sup><http://www.ipac.caltech.edu/H0kp/>

<sup>14</sup>In fact, one might even consider a simpler case with only 5 free parameters by fixing  $n_s = 1$  (Liddle 2004).

surement of the acoustic scale from the SDSS LRG power spectrum as given in Hütsi (2006a). The analysis in the second part uses the full power spectrum measurement along with the covariance matrix as provided by Hütsi (2006a). Here we add two extra parameters: bias parameter  $b$  and parameter  $Q$  that describes the deformation of the linear power spectrum to the nonlinear redshift-space spectrum. These extra parameters are treated as nuisance parameters and are marginalized over. Thus the largest parameter space we should cope with is 9-dimensional.<sup>15</sup> Since the parameter space is relatively high dimensional it is natural to use Markov Chain Monte Carlo (MCMC) techniques. For this purpose we use publicly available cosmological MCMC engine Cosmomc<sup>16</sup> (Lewis & Bridle 2002). All the CMB spectra and matter transfer functions are calculated using the fast Boltzmann code CAMB<sup>17</sup> (Lewis et al. 2000).

The paper is organized as follows. In Sect. 2 we describe the observational data used for the parameter estimation. Sect. 3 discusses and tests the accuracy of the transformations needed to convert the linear input spectrum to the observed redshift-space galaxy power spectrum. In Sect. 4 we present the main results of the cosmological parameter study and we conclude in Sect. 5.

## 4.2. Data

The SDSS LRG power spectrum as determined by Hütsi (2006a) is shown with filled circles and heavy solid errorbars in Fig. 4.1. There the upper data points correspond to the deconvolved version of the spectrum.<sup>18</sup> The thin solid lines represent the best-fitting model spectra, with the lower curve corresponding to the convolved case. As the survey window is relatively narrow the deconvolution can be done rather “cleanly”. This deconvolved spectrum might be useful for the extra-fast parameter estimation employing analytic approximations for the matter transfer functions (Eisenstein & Hu 1998, Novosyadlyj et al. 1999) and fast CMB angular power spectrum generators such as CMBfit<sup>19</sup> (Sandvik et al. 2004), DASH<sup>20</sup> (Kaplinghat et al. 2002) and CMBwarp<sup>21</sup> (Jimenez et al. 2004). However, in this paper, as we use an accurate Boltzmann solver CAMB to calculate CMB power spectra and matter transfer functions, the relative time taken by an extra convolution step is completely negligible. Thus in the following we use only the convolved spectrum.<sup>22</sup> Accurate analytic fitting formulae for the survey window functions can be found in Hütsi (2006a).<sup>23</sup> The power spectrum covariance matrix in Hütsi (2006a) was measured from 1000 mock catalogs generated with the second-order optimized Lagrangian perturbation calculation. The same paper also provides the measurement of the acoustic scale:  $(105.4 \pm 2.3) h^{-1}$  Mpc. This corresponds to the case when only sinusoidal modulation, as expected in the case of adiabatic initial conditions, in the power spectrum is allowed. Relaxing this assumption by allowing

<sup>15</sup>Since marginalization over the bias parameter can be done analytically (Bridle et al. 2002) the actual number of parameters can be reduced to 8.

<sup>16</sup><http://cosmologist.info/cosmomc/>

<sup>17</sup><http://camb.info/>

<sup>18</sup>The deconvolution was performed using an iterative algorithm due to Lucy (1974) with a specific implementation as given in Lin et al. (1996).

<sup>19</sup><http://www.hep.upenn.edu/sandvik/CMBfit.html>

<sup>20</sup><http://bubba.ucdavis.edu/DASH/>

<sup>21</sup><http://www.physics.upenn.edu/raulj/CMBwarp/>

<sup>22</sup>Often also called a pseudospectrum.

<sup>23</sup>There the combination ‘mode coupling kernels’ is used in place of the more common ‘window functions’.

an arbitrary phase shifts gave the result,  $(103.0 \pm 7.6) h^{-1} \text{ Mpc}$ , instead. In the following parameter estimation process we use both of these values. In [Hütsi \(2006a\)](#) the measurement of the acoustic scale was achieved by first removing the “smooth” component of the spectrum and then fitting the parametrized family of functions to the oscillatory part via the modified version of the Levenberg-Marquardt method. The separation of the “smooth” and “oscillatory” components of the spectrum can be done rather accurately since the characteristic scales over which they change differ strongly. The Levenberg-Marquardt method which was used to determine the oscillation frequency approximates the likelihood surface near its maximum with a multidimensional Gaussian, and this way provides an approximate parameter covariance matrix. To avoid this “Gaussianity assumption” we have also performed a MCMC parameter estimation exercise, finding the best fitting acoustic scale along with its uncertainty in full agreement with the values quoted above. The question that might arise of course is how adequate is the parametric family that was used for fitting the oscillatory component? Even in the simplest case of the adiabatic initial fluctuations the damped sinusoidal modulation is only an approximation. We investigate the possible biases introduced by assuming a fixed parametric form for the oscillatory part of the spectrum in more detail in Sect. 4.4.2.

As mentioned in the Introduction, in order to break several degeneracies between the cosmological parameters, we complement the SDSS LRG power spectrum data with the data from the WMAP CMB measurements. Specifically, we use the CMB temperature power spectrum as found in [Hinshaw et al. \(2003\)](#) and the temperature-polarization cross-power as determined by [Kogut et al. \(2003\)](#). The description of the likelihood calculation using this data is given in [Verde et al. \(2003\)](#). We use the Fortran90 version of this likelihood code as provided by the Cosmomc package.

While investigating the constraints arising from the measurement of the acoustic scale we do not run each time the full new MCMC calculation. Instead we importance sample the chains built for the WMAP data along with the constraint on the Hubble parameter as provided by the HST Key Project,  $H_0 = 72 \pm 8 \text{ km/s/Mpc}$  ([Freedman et al. 2001](#)). Using the WMAP data alone would result in too loose constraints on several parameters, and thus after importance sampling a large fraction of the chain elements would get negligible statistical weight, leaving us with too small effective number of samples.

### 4.3. Power spectrum / acoustic scale transformation

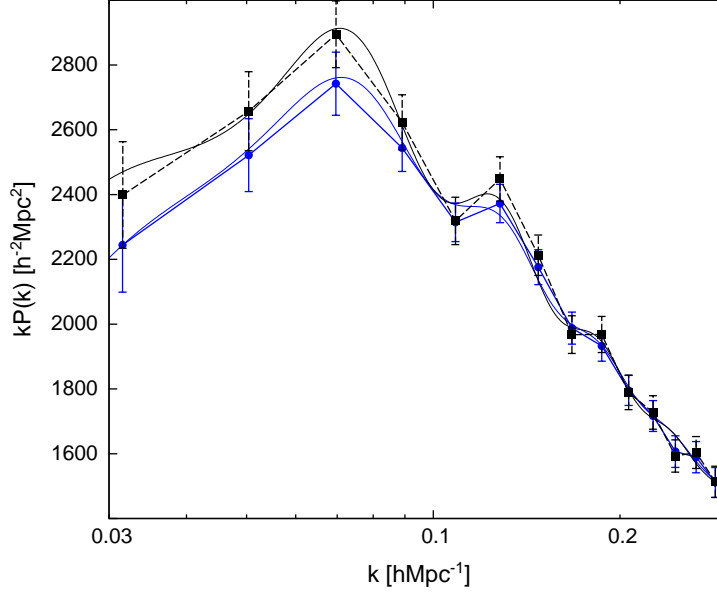
In this section we discuss the relation of the observed galaxy power spectrum to the underlying spectrum of the matter distribution. We stress the need to take into account the so-called cosmological distortion, which almost always is being completely neglected.<sup>24</sup> This is fine for the very shallow surveys, but as we show later, for the samples like the SDSS LRGs, with an effective depth of  $z_{\text{eff}} \sim 0.35$ , the cosmological distortion should certainly be taken into account. This is especially important if power spectrum, instead of being well approximated by a simple power law, contains some characteristic features.

There are other difficulties one has to face while trying to make cosmological inferences using the observed galaxy samples. It is well known that galaxies need not faithfully follow the

---

<sup>24</sup>According to our knowledge the only counter-example being the work by [Eisenstein et al. \(2005\)](#).



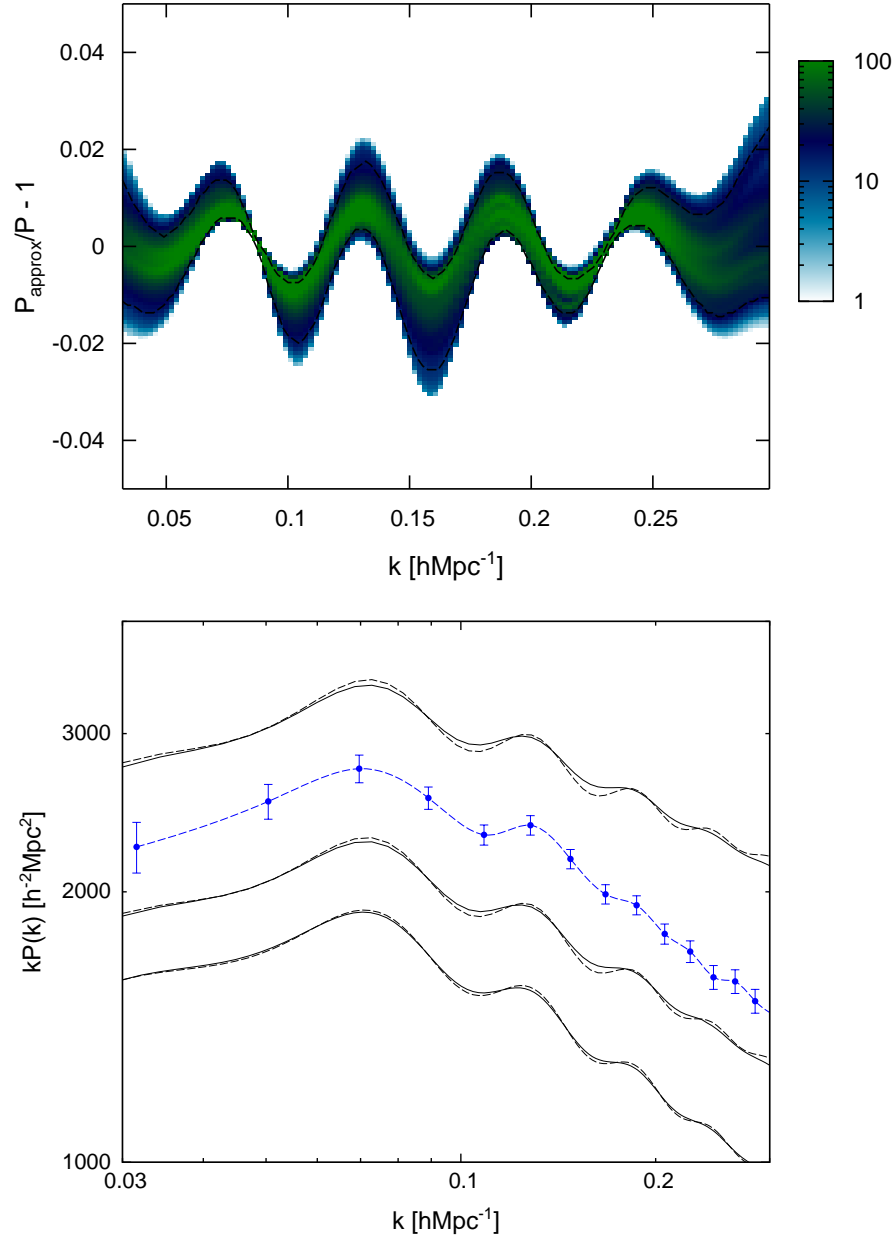


**Figure 4.1.:** Filled circles with solid errorbars represent the SDSS LRG power spectrum as determined by [Hütsi \(2006a\)](#). The upper data points provide the deconvolved version of the spectrum. The thin solid lines show the best-fitting model spectra.

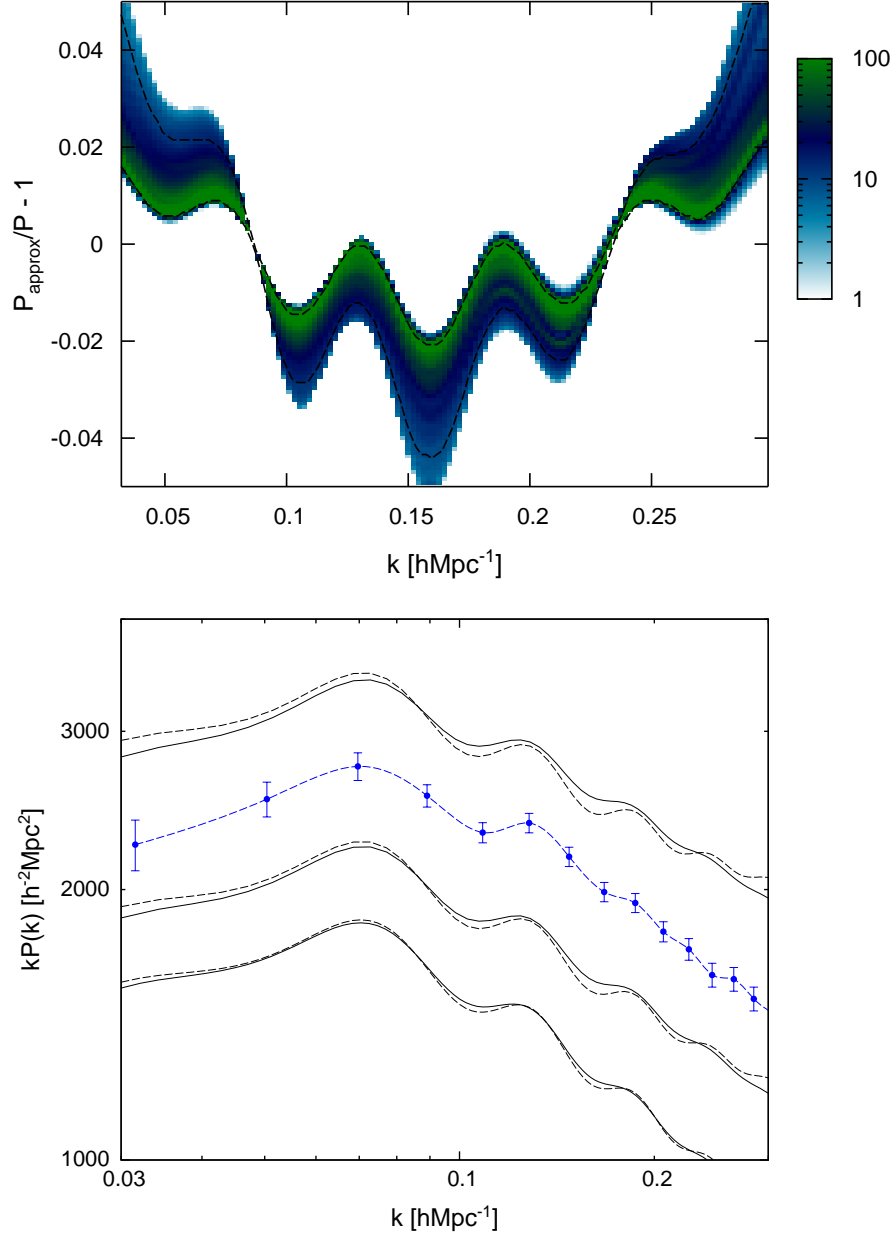
underlying matter distribution. This phenomenon is known as biasing ([Kaiser 1984](#)). Whereas on the largest scales one might expect linear and scale-independent biasing (e.g. [Coles 1993](#), [Narayanan et al. 2000](#)), on smaller scales this is definitely not the case. In general the biasing can be scale-dependent, nonlinear, and stochastic ([Dekel & Lahav 1999](#)). The other complications involved are the redshift-space distortions and the effects due to nonlinear evolution of the density field. The redshift-space distortions, biasing, and nonlinearities can be approximately treated in the framework of the Halo Model approach as described in Appendix D. The implementation of the Halo Model as presented there introduces four new parameters:  $M_{\text{low}}$ , the lower cutoff of the halo mass i.e. below that mass halos are assumed to be “dark”;  $\alpha$  and  $M_0$ , the parameters describing the mean of the halo occupation distribution i.e. the average number of galaxies per halo with mass  $M$ , which was assumed to have the form  $\langle N|M \rangle = \left(\frac{M}{M_0}\right)^\alpha$ ;  $\gamma$ , the parameter describing the amplitude of the virial motions inside the haloes with respect to the isothermal sphere model. This formulation of the Halo Model, along with the assumption of the best-fit WMAP cosmology ([Spergel et al. 2003](#)), is able to produce a very good fit to the observed SDSS LRG power spectrum as demonstrated in Fig. 4.1. Moreover, all the parameters:  $M_{\text{low}}$ ,  $\alpha$ ,  $M_0$ ,  $\gamma$ , are reasonably well determined. It turns out that to a good approximation these four extra parameters can be compressed down just to a single parameter  $Q$ , describing the deformation of the linearly evolved spectrum:

$$P_{\text{gal}}(k) = b^2(1 + Qk^\eta)P_{\text{lin}}(k). \quad (4.1)$$

Here  $b$  is the bias parameter and a good value for  $\eta$  turns out to be  $\frac{3}{2}$ . A similar type of parametric description for the galaxy power spectrum was also used in [Cole et al. \(2005\)](#), with a slight difference for the treatment of the largest scales. In that paper the authors suggest to take



**Figure 4.2.:** Upper panel: A density plot showing the probability distribution functions for the relative accuracy of the approximation given in Eq. (4.1) with  $\eta = \frac{3}{2}$ . The set of Halo Model parameters  $M_{\text{low}}$ ,  $\alpha$ ,  $M_0$ , and  $\gamma$ , needed to calculate the “exact” spectra, were drawn from the multidimensional Gaussian distribution centered at the best-fit values and with a covariance matrix as found in [Hütsi \(2006a\)](#). The heavy dashed lines mark the 5% and 95% quantiles of the relative accuracy distributions. Lower panel: Filled circles with solid errorbars provide the SDSS LRG power spectrum. The data points are connected with a smooth cubic spline fit. The other set of lines represents some examples of the pairs of spectra that correspond, starting from below, to the best matching case, to the 68%, and to the 90% quantiles of the distribution of the  $\chi^2$  values. The solid lines show the Halo Model spectra while the dashed ones are the approximations from Eq. (4.1).



**Figure 4.3.:** The version of Fig. 4.2 with  $\eta = 2$ .

$\eta = 2$ . However, we have found that  $\eta = \frac{3}{2}$  provides a better approximation for these 4-parameter Halo Model spectra.<sup>25</sup> This is demonstrated in the upper panels of Figs. 4.2 and 4.3. There we have calculated a full range of Halo Model spectra (assuming the WMAP “concordance” cosmology) for different values of  $M_{\text{low}}$ ,  $\alpha$ ,  $M_0$ , and  $\gamma$ , drawn from the multidimensional Gaussian centered around the best-fitting values and with the parameter covariance matrix as found in Htsi (2006a). Each of the calculated models is fitted with a simple parametric form as given in Eq. (4.1). The upper panels of Figs. 4.2 and 4.3 represent the density plots for the fractional accuracy of these simple fits i.e. for each wavenumber  $k$  they show the probability distribution functions for the achieved relative accuracy. With the heavy dashed lines we have also marked the 5% and 95% quantiles of the accuracy distributions. It is evident that  $\eta = \frac{3}{2}$  provides significantly better approximation to the spectral deformation than  $\eta = 2$ . The largest errors are seen to be located at the positions of the acoustic features, with a simple approximation in Eq. (4.1) giving larger oscillation amplitudes. The Halo Model gives lower oscillation amplitudes since relatively flat contribution from the 1-halo term added to the 2-halo part starts to decrease the contrast of the acoustic features, whereas the multiplicative transform in Eq. (4.1) preserves the contrast level of these wiggles. In the lower panels of Figs. 4.2 and 4.3 we have provided some examples of the pairs of spectra that correspond (staring from below) to the best matching case, and also the ones representing the 68% and 90% quantiles of the distribution of the  $\chi^2$  values. The solid lines here correspond to the Halo Model calculations. For comparison also the SDSS LRG power spectrum along with the cubic spline fitted to the data points are shown. For clarity slight vertical shifts have been applied to the model spectra. As can be seen, the approximation in Eq. (4.1) is rather acceptable in the light of the accuracy of the SDSS LRG power spectrum measurement. This approximation is used in Sect. 4.4.3 where we fit the model spectra to the SDSS LRG data.

The cosmological distortion, mentioned in the beginning of this section, arises due to the simple fact that conversion of the observed redshifts to comoving distances requires the specification of the cosmological model. If this cosmology differs from the true one, we are left with additional distortion of distances along and perpendicular to the line of sight. In general, the spatial power spectrum measurements, in contrast to the angular spectra, are model dependent i.e. along with the measurements of the 3D power spectrum one always has to specify the so-called fiducial model used to analyze the data. The fiducial model corresponding to the data shown in Fig. 4.1 is the best-fit WMAP “concordance” model (Spergel et al. 2003). In principle, for each of the fitted cosmological model one should redo the full power spectrum analysis to accommodate different distance-redshift relation. However, there is an easier way around: one can find an approximate analytical transformation that describes how the model spectrum should look like under the distance-redshift relation given by the fiducial model i.e. instead of transforming the data points we transform the fitted model spectra. Since the distance intervals along and perpendicular to the line of sight transform differently, the initial isotropic theoretical spectrum  $P$  transforms to the 2D spectrum:

$$\tilde{P}^{2D}(k_{\parallel}, k_{\perp}; z) = \frac{1}{c_{\parallel}(z) \cdot c_{\perp}^2(z)} P \left[ \sqrt{\left( \frac{k_{\parallel}}{c_{\parallel}(z)} \right)^2 + \left( \frac{k_{\perp}}{c_{\perp}(z)} \right)^2}; z \right], \quad (4.2)$$

<sup>25</sup>At least if the spectra have shapes close to the observed SDSS LRG spectrum.

where the distortion parameters along and perpendicular to the line of sight are defined as:

$$c_{\parallel}(z) = \frac{H^{\text{fid}}(z)}{H(z)}, \quad (4.3)$$

$$c_{\perp}(z) = \frac{d_{\perp}(z)}{d_{\perp}^{\text{fid}}(z)}. \quad (4.4)$$

Here  $H(z)$  is the Hubble parameter and  $d_{\perp}(z)$  is the comoving angular diameter distance corresponding to the fitted theoretical model. Superscript  $\text{fid}$  refers to the fiducial model. Here and in the following we use a tilde on top of  $P$  to denote theoretical spectrum “transformed to the reference frame of the fiducial cosmology”. As we use the spectra that have the dimensions of volume an extra division by  $c_{\parallel}(z) \cdot c_{\perp}^2(z)$  occurs due to the transformation of the volume elements:

$$dV(z) = c_{\parallel}(z) \cdot c_{\perp}^2(z) \cdot dV^{\text{fid}}(z). \quad (4.5)$$

By introducing the variables

$$k = \sqrt{k_{\parallel}^2 + k_{\perp}^2}, \quad \mu = \frac{k_{\parallel}}{k}, \quad \kappa(z) = \frac{c_{\parallel}(z)}{c_{\perp}(z)}, \quad (4.6)$$

we can express  $\tilde{P}^{2D}$  as follows:

$$\tilde{P}^{2D}(k, \mu; z) = \frac{1}{c_{\parallel}(z) \cdot c_{\perp}^2(z)} P \left[ \frac{k}{c_{\perp}(z)} \sqrt{1 + \left( \frac{1}{\kappa^2(z) - 1} \right) \mu^2}; z \right]. \quad (4.7)$$

Now the corresponding isotropized spectrum can be given as:

$$\tilde{P}(k; z) = \frac{1}{2 c_{\parallel}(z) \cdot c_{\perp}^2(z)} \int_{-1}^1 P \left[ \frac{k}{c_{\perp}(z)} \sqrt{1 + \left( \frac{1}{\kappa^2(z) - 1} \right) \mu^2}; z \right] d\mu. \quad (4.8)$$

As the observations are done along the light-cone we have to perform relevant integrals along the redshift. The full treatment for the light-cone effect can be found in [Yamamoto & Suto \(1999\)](#), [Yamamoto et al. \(1999\)](#). As we are investigating a two-point function, an accurate light-cone calculation would introduce two integrals over the redshifts ([Matarrese et al. 1997](#), [Yamamoto et al. 1999](#)). However, it turns out that to a good approximation, excluding the very largest scales, the contributions from different redshifts decouple and the double integral reduces to a simple one-dimensional integral over redshift. The final result for the  $\tilde{P}(k; z)$ , averaged over the light-cone can be given as:

$$\tilde{P}(k) = \frac{\int_{z_{\min}}^{z_{\max}} \frac{dV^{\text{fid}}}{dz} dz \cdot \mathcal{W}^2(k; z) \bar{n}^2(z) \tilde{P}(k; z) c_{\parallel}(z) c_{\perp}^2(z)}{\int_{z_{\min}}^{z_{\max}} \frac{dV^{\text{fid}}}{dz} dz \cdot \mathcal{W}^2(k; z) \bar{n}^2(z) c_{\parallel}(z) c_{\perp}^2(z)}. \quad (4.9)$$

Here the result of Yamamoto et al. (1999) has been generalized to include other weight factors in addition to the simple number density weighting. The most common weight functions  $\mathcal{W}(z)$  are the following:

$$\mathcal{W}(k; z) \propto \begin{cases} \frac{1}{\bar{n}(z)} & \text{for volume weighting} \\ \text{const} & \text{for number weighting} \\ \frac{1}{1 + \bar{n}(z)\tilde{P}(k; z)} & \text{for the FKP weighting.} \end{cases} \quad (4.10)$$

Here FKP stands for the weighting scheme due to Feldman et al. (1994). The power spectrum measurement of the SDSS LRGs in (Hütsi 2006a) used the FKP weighting function. In Eqs. (4.9) and (4.10)  $\bar{n}(z)$  represents the mean number density of galaxies as a function of redshift. For the SDSS LRG sample analyzed in Hütsi (2006a) the limiting redshifts  $z_{\min} = 0.16$  and  $z_{\max} = 0.47$ . If instead of the integral over  $z$  in Eq. (4.9) we just take the integrand at the effective redshift (e.g. the median redshift) of the survey, and replace the distortion parameters  $c_{\parallel}$  and  $c_{\perp}$  with a single “isotropized” dilation of scales (see e.g. Eisenstein et al. 2005):

$$c_{\text{isotr}} = \sqrt[3]{c_{\parallel}(z_{\text{eff}}) c_{\perp}^2(z_{\text{eff}})}, \quad (4.11)$$

we can write instead of Eq. (4.9) simply

$$\tilde{P}(k) = \frac{1}{c_{\text{isotr}}} P\left(\frac{k}{c_{\text{isotr}}}\right). \quad (4.12)$$

Here the prefactor  $1/c_{\text{isotr}}$  can also be dropped, as it can be absorbed into the bias parameter that is assumed to be a completely free parameter throughout this paper. Although the true transformation for the power spectrum is different along and perpendicular to the line of sight, and also is dependent on redshift, it turns out that a single dilation approximation taken at the median redshift of the survey can provide a very good approximation, especially for relatively shallow surveys. For the median redshift of the SDSS LRG sample as analyzed in Hütsi (2006a),  $z \sim 0.35$ , this approximation is very accurate as can be seen in Fig. 4.4. The upper panel of Fig. 4.4 shows a similar density plot as in Figs. 4.2 and 4.3. Here, in comparison to Figs. 4.2 and 4.3 where the background cosmology was fixed to the best-fit WMAP model and the Halo Model parameters were varied, we use the simple linear spectra while changing the cosmology. The set of cosmological models is drawn from the combined posterior corresponding to the WMAP plus HST Key Project data. As can be seen from this figure, for relatively shallow surveys the single “isotropized” dilation approximation is very precise: for  $\sim 90\%$  of the models the approximation in Eq. (4.12) is more accurate than 0.5%. This is even more clear when looking at the lower panel of Fig. 4.4 where we have plotted the pairs of spectra corresponding to the best matching case, and also some examples representing 68% and 90% quantiles of the distribution of the  $\chi^2$  values. As can be seen, even the pair of curves corresponding to the 90% quantile, are basically indistinguishable. In Fig. 4.5 we have illustrated the case when the cosmological distortion is ignored. One can see that for  $\sim 90\%$  of cases we make relative errors of  $\sim 6\%$ , which is comparable to the amplitude of the acoustic oscillatory features. The lower panel of Fig. 4.5 presents pairs of spectra for 68% and 90% quantiles of the  $\chi^2$  values. The inset shows the probability distribution function for the “isotropized” dilation scale, as given in Eq. (4.11), compatible with the WMAP plus HST Key project constraints. Since the values of  $c_{\text{isotr}}$  are quite

often seen to differ from  $c_{\text{isotr}} = 1$  by 5 – 10%, it is clear that the cosmological distortion has to be taken into account if the power spectrum is measured as accurately as given by the SDSS LRG data points in the lower panel of Fig. 4.5.

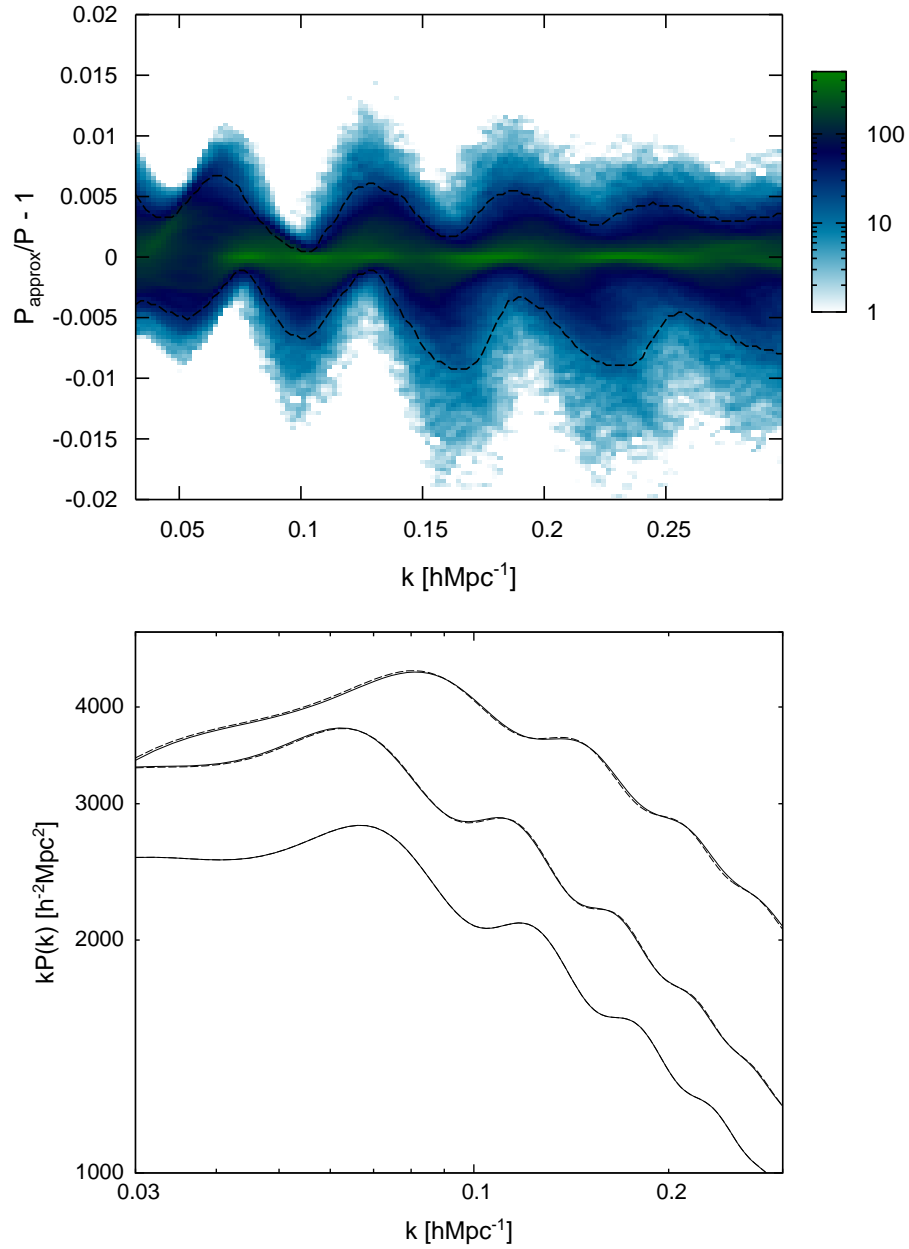
## 4.4. Results

### 4.4.1. $W_{\text{MAP}}$ + HST data

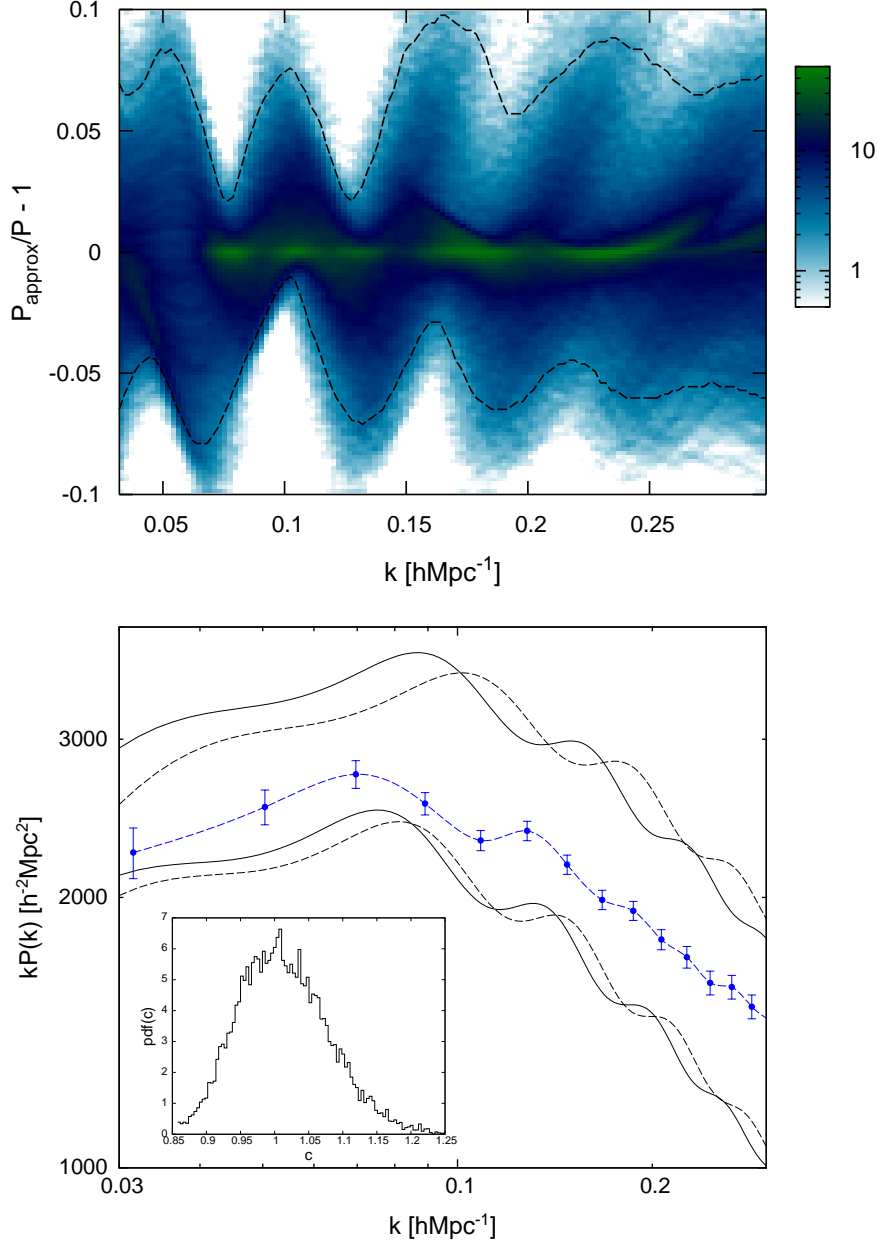
As a starting point for several subsequent calculations we build a Markov chain using the  $W_{\text{MAP}}$  temperature-temperature (Hinshaw et al. 2003) and temperature-polarization (Kogut et al. 2003) angular spectra in combination with the constraint on the Hubble parameter from the HST Key Project (Freedman et al. 2001). The results for the 2D marginalized distributions for all of the involved parameter pairs are shown in Fig. 4.6. Here the 68% and 95% credible regions are shown by solid lines. The original MCMC calculation as performed by the Cosmomc software uses the variable  $\theta$  – the angle subtended by the sound horizon at last scattering – in place of the more common Hubble parameter  $H_0$ . This leads to the better mixing of the resulting chain since  $\theta$  is only weakly correlated with other variables (Kosowsky et al. 2002). The proposal distribution for all of the MCMC calculations carried out in this paper is taken to be a multivariate Gaussian. For the current  $W_{\text{MAP}}$  + HST case we have used the CMB parameter covariance matrix as provided by the Cosmomc package. All of the seven default parameters here get implicit flat priors. The marginalized distributions in Fig. 4.6 are derived from a 100,000-element Markov chain. As there is a very good proposal distribution available the chains typically equilibrate very fast and only a few hundred first elements need to be removed to eliminate the effects of the initial transients. We determine the length of this so-called burn-in period using the Gibbsit<sup>26</sup> software (Raftery & Lewis 1995). The same program can also be used to estimate the length of the Markov chain required to achieve a desired accuracy for the parameter measurements. As a test one can run initially a short chain of a few thousand elements and analyze it with Gibbsit. It turned out that in the current case if we would like to achieve a 1.25% accuracy at 95% confidence level for the measurement of the 2.5% and 97.5%-quantiles of the most poorly sampled parameter, we would need a chain of  $\sim 25,000$  elements. Thus according to this result our 100,000 element chain is certainly more than sufficient. Of course, all the various tools for diagnosing the convergence and for estimating the required chain length<sup>27</sup> are just some more or less justified “recipes” that can lead to strongly incorrect results, especially in cases of poorly designed proposal distributions. Luckily, in cosmology as we have a very good knowledge about the possible parameter degeneracies, and also as the parameter spaces are relatively low dimensional, the construction of very good samplers is not too difficult.

In the following subsections we use this  $W_{\text{MAP}}$  + HST chain for the very fast determination of the parameter constraints resulting from the additional measurement of the SDSS LRG acoustic scale. The same chain was also used to produce Figs. 4.4 and 4.5.

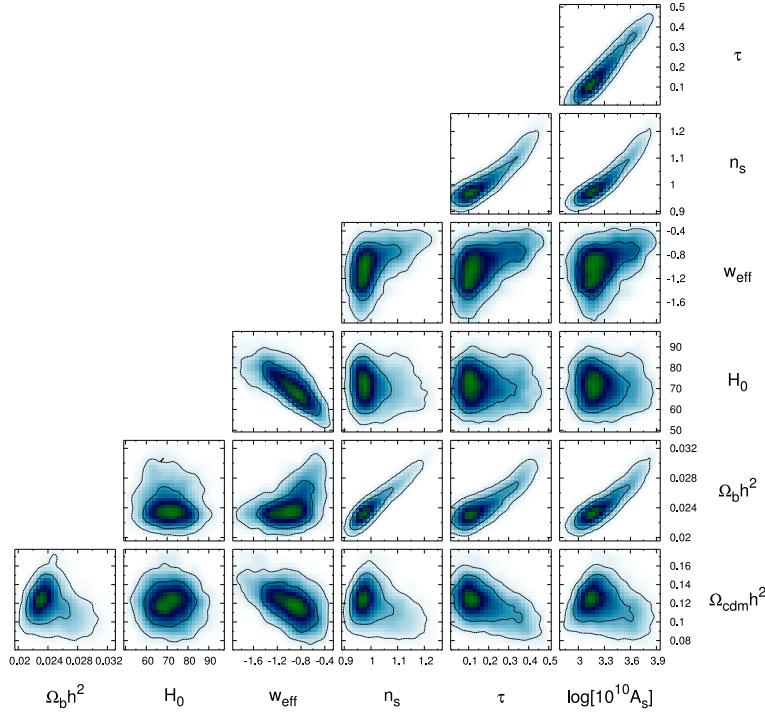




**Figure 4.4.:** An analog of Fig 4.2, here provided in the context of the accuracy test for the cosmological distortion approximation given in Eq. (4.12). The set of cosmological models was drawn from the combined posterior corresponding to the  $W_{\text{MAP}}$  plus HST Key Project data.



**Figure 4.5.:** As Fig. 4.4, here instead showing the error one makes if cosmological distortion is completely neglected. In the lower panel we have shown only the examples corresponding to the 68% and 90% quantiles. The inset shows the probability distribution function for the “isotropized” dilation scale, as given in Eq. (4.11), compatible with the  $W_{\text{MAP}}$  plus HST Key project constraints.



**Figure 4.6.:** The 2D marginalized distributions for the WMAP + HST data.

#### 4.4.2. Constraints from the measurement of the acoustic scale

The low redshift acoustic scale as measured via the analysis of the SDSS LRG power spectrum was found to be  $(105.4 \pm 2.3) h^{-1} \text{ Mpc}$  if adiabatic initial conditions were assumed (i.e. allowing only for the sinusoidal modulation in the spectrum), and  $(103.0 \pm 7.6) h^{-1} \text{ Mpc}$  if this assumption was relaxed by allowing additional oscillation phase shifts (Hütsi 2006a). These measurements refer to the WMAP best-fit cosmology (Spergel et al. 2003) which was used to analyze the SDSS LRG data. In Sect. 4.3 we have described accurate transformations needed to accommodate other background cosmologies. In the following we use SH1 and SH2 to denote the sound horizon measurements of  $(105.4 \pm 2.3) h^{-1} \text{ Mpc}$  and  $(103.0 \pm 7.6) h^{-1} \text{ Mpc}$ , respectively.

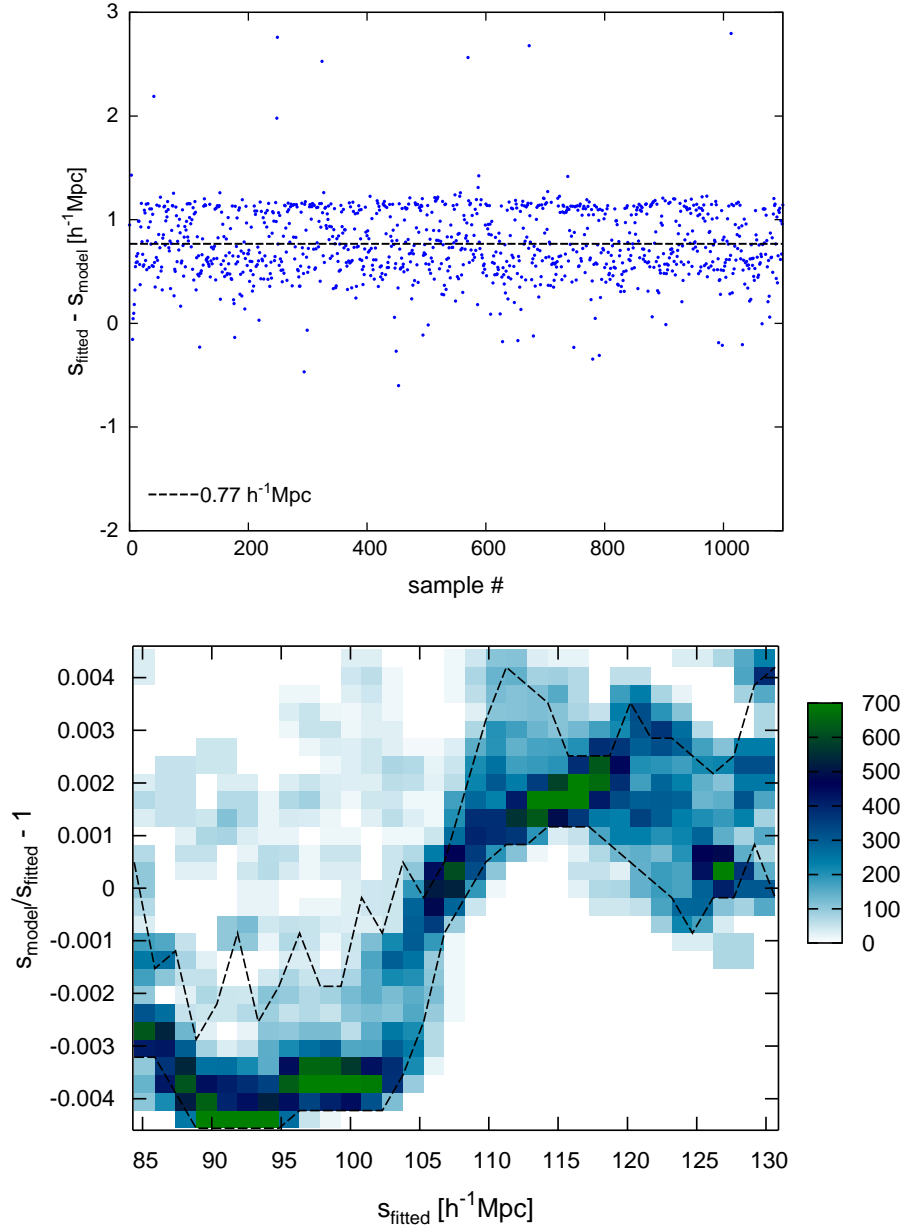
In this section we investigate the constraints on cosmological parameters using the above given values for the sound horizon in combination with the WMAP data. To speed up the calculations we do not build new Markov chains from scratch, instead we apply the method of importance sampling on the earlier calculated WMAP + HST chain. It is fine to use importance sampling if new constraints are not too constraining and are consistent with the earlier generated chain. Having a measurement of the acoustic scale  $\tilde{s}$ <sup>28</sup> with an error  $\Delta\tilde{s}$ , importance sampling simply amounts to multiplying each original sample weight by

$$f_i = \exp \left[ -\frac{(\tilde{s}_{\text{model}_i} - \tilde{s})^2}{2 \Delta\tilde{s}^2} \right], \quad (4.13)$$

<sup>26</sup><http://www.stat.washington.edu/raftery/software.html>

<sup>27</sup>For a lot of online material related to these issues see <http://www.statslab.cam.ac.uk/~mcmc/>.

<sup>28</sup>We use tilde to denote the quantities that are “tied to the” fiducial cosmological model used to analyze the data.



**Figure 4.7.:** Upper panel: Comparison of the sound horizon as determined from  $\sim 1000$  model spectra, via the same fitting techniques that were used in [Hütsi \(2006a\)](#) to measure the SDSS LRG sound horizon, with the analytical approximation given in Eqs. (A.4), (A.5), (A.9), (A.10), (A.11). The model spectra were drawn from the posterior distribution corresponding to the WMAP + HST data. Lower panel: The density plot of the residuals after removing the average bias of  $0.77 h^{-1} \text{Mpc}$ . The solid dashed lines mark the 68% credible region.

where  $\tilde{s}_{\text{model}_i}$  denotes the theoretical sound horizon corresponding to the  $i$ -th Markov chain element. The physical size of the sound horizon  $s$  at the end of the drag-epoch is determined by the parameter combinations  $\Omega_m h^2$  and  $\Omega_b h^2$  i.e. physical densities of the CDM and baryonic components. Accurate fitting formulae for  $s$  can be found in [Hu & Sugiyama \(1996\)](#), [Eisenstein & Hu \(1998\)](#). We have provided these in Appendix A where also the transformation into different cosmological frame is described. This transformation induces an extra dependence of the sound horizon  $\tilde{s}$ , as measured from the matter power spectrum, on  $h$  and  $w_{\text{eff}}$ . For more details see Appendix A. The dependence of  $\tilde{s}$  at redshift  $z \sim 0.35$  on various parameters for spatially flat models around the best fitting WMAP model point can be conveniently expressed as the following principal component:

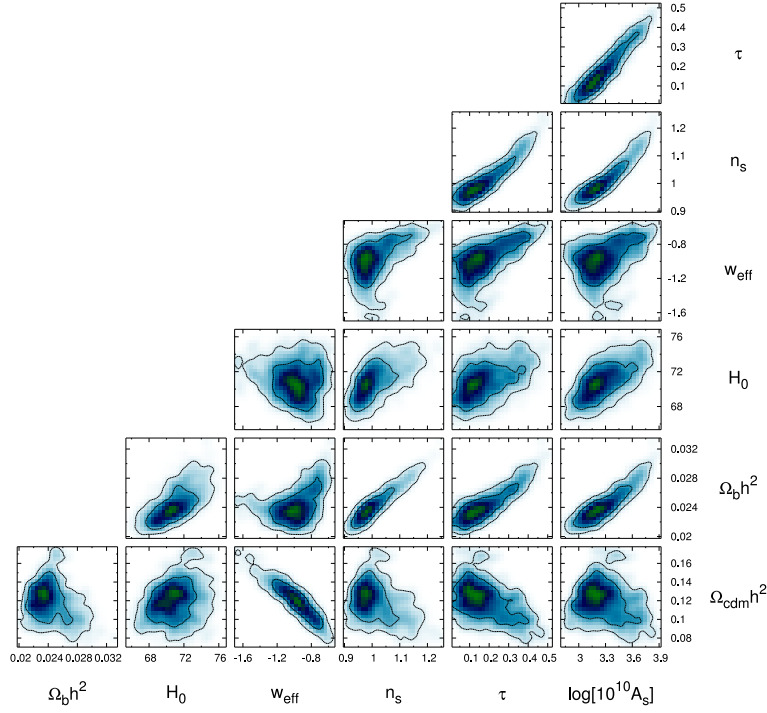
$$\left(\frac{\Omega_m h^2}{0.14}\right)^{-0.28} \left(\frac{\Omega_b h^2}{0.022}\right)^{-0.10} \left(\frac{h}{0.71}\right)^{0.94} \left(\frac{w_{\text{eff}}}{-1.}\right)^{0.14} = 1 \pm \frac{\Delta \tilde{s}}{\tilde{s}}. \quad (4.14)$$

As probably expected, for those relatively small redshifts by far the strongest dependence is on the Hubble parameter  $h$ . To avoid any biases due to the approximate nature of the Eqs. (A.9), (A.10), (A.11), and also due to the specific method used to measure the oscillation frequency in the SDSS LRG power spectrum, we carry out the following Monte Carlo study. We draw  $\sim 1000$  samples from the WMAP + HST chain by thinning it by a factor of  $\sim 10$ . For each of the parameter combinations we calculate theoretical matter spectra using CAMB. The oscillatory components of the spectra are extracted by dividing them with a “smoothed” approximate model spectra as given in [Eisenstein & Hu \(1998\)](#).<sup>29</sup> The resulting “flattened” spectra are fitted with damped sinusoidal waves<sup>30</sup> and the sound horizon  $\tilde{s}_{\text{fitted}}$  is determined via the Levenberg-Marquardt fitting technique. All the spectra are calculated at exactly the same wavenumbers as the data points given in Fig 4.1. The power spectrum covariance matrix is taken from the Appendix G of [Hütsi \(2006a\)](#). For each model the sound horizon  $\tilde{s}_{\text{model}}$  is calculated using Eqs. (A.4), (A.5), (A.9), (A.10), (A.11). The comparison of  $\tilde{s}_{\text{fitted}}$  versus  $\tilde{s}_{\text{model}}$  is provided in Fig. 4.7. In the upper panel we have plotted  $\tilde{s}_{\text{fitted}} - \tilde{s}_{\text{model}}$ . As can be seen there is a slight tendency for the fitted values  $\tilde{s}_{\text{fitted}}$  to be larger than  $\tilde{s}_{\text{model}}$ . After removing the constant bias of  $0.77 h^{-1} \text{ Mpc}$  the remaining fluctuations are  $\lesssim 0.5\%$ , which is demonstrated in the lower panel of Fig. 4.7. This plot is an analog to the earlier density plots shown in Figs. 4.2, 4.3, 4.4, 4.5. Here the dashed lines show the region covering 68% of all the cases. One might even go a step further and instead of removing only a constant offset, remove also the next order i.e. the linear component. This more accurate treatment has probably rather negligible effect on the final results, since around the measured sound horizon values of  $\sim 105 h^{-1} \text{ Mpc}$  the accuracy after removing the constant offset is already  $\sim 0.2 - 0.3\%$ , which is an order of magnitude smaller than the measurement errors of  $2.3 - 7.6 h^{-1} \text{ Mpc}$ .

Using this correction for the bias and the method to calculate the theoretical size of the sound horizon at the end of the drag-epoch, as presented in Appendix A, we can immediately perform the relevant reweighting of the WMAP + HST chain (see Eq. (4.13)). It turns out that relatively large fraction of the WMAP + HST chain elements “survive” this reweighting procedure, justify-

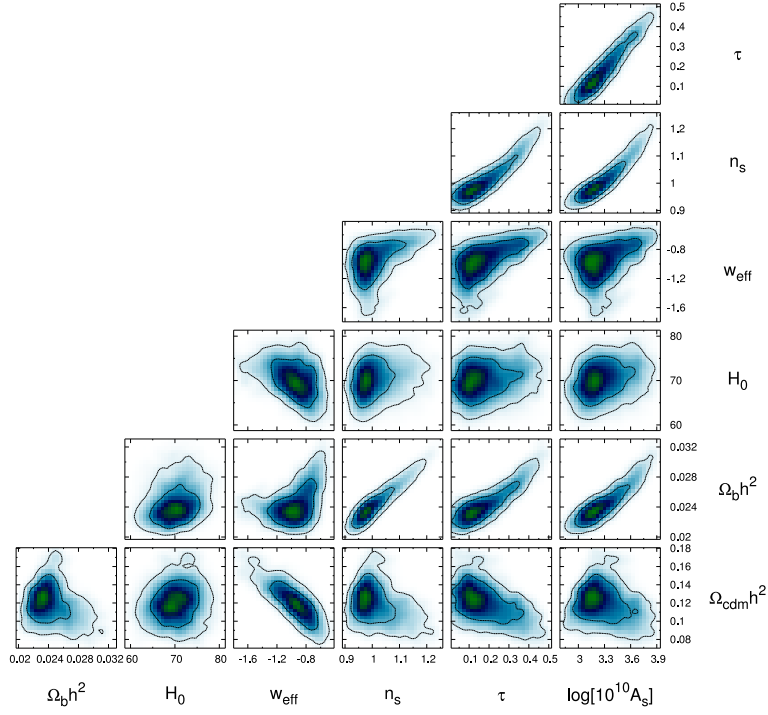
<sup>29</sup>The separation of the oscillatory component and the underlying smooth CDM continuum can be done very cleanly due to significantly different characteristic scales over which they change. The small residual deformations of the oscillatory part have negligible impact on the inferred oscillation period.

<sup>30</sup>For a precise parametric form see [Hütsi \(2006a\)](#).

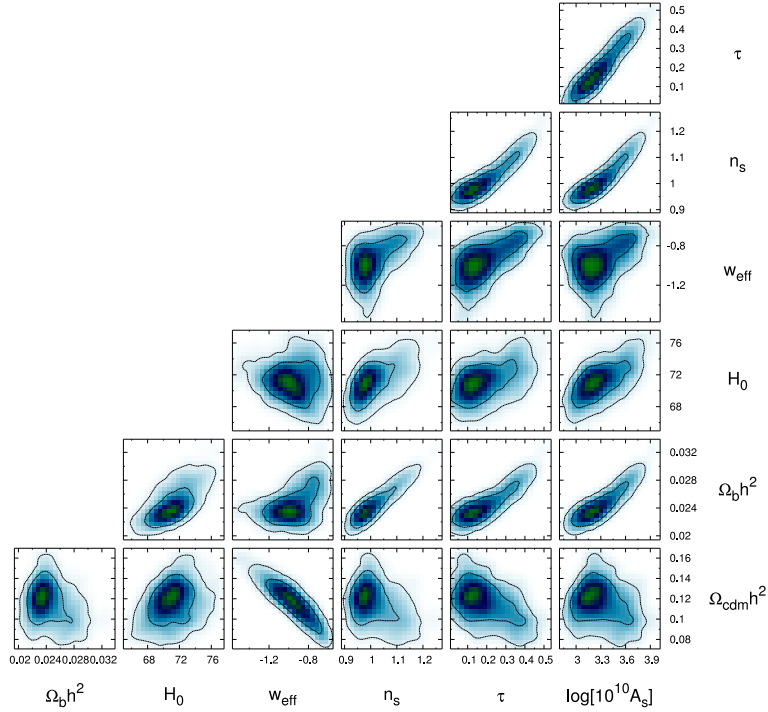


**Figure 4.8.:** The 2D marginalized distributions for the WMAP data along with the constraint on the low redshift sound horizon,  $\widetilde{r}(105.4 \pm 2.3) h^{-1}$  Mpc (SH1), obtained via the importance sampling of the WMAP + HST results shown in Fig. 4.6.

ing the use of the importance sampling method. In particular, for the SH2 case we are left with  $\sim 36,000$ , and for the SH1  $\sim 12,000$  samples. The results of this calculation in the form of the 2D marginalized distributions is presented in Figs. 4.8 and 4.9. Here Fig. 4.8/4.9 corresponds to the SH1/SH2 case. In comparison to the analogous Fig. 4.6 the most dramatic changes are for  $H_0$  and  $w_{\text{eff}}$ , whereas the rest of the parameters stay essentially the same. The HST constraint for the initial WMAP chain was just implemented in order not to loose too many samples in current importance sampling calculations. As can be seen, the new constraints on  $H_0$  are significantly stronger than the one provided by the HST. In Fig. 4.8 due to somewhat lower number of samples ( $\sim 12,000$ ) the contours start to become more noisy. Earlier we estimated that the measurement of the 2.5% and 97.5% quantiles with an accuracy of 1.25% at 95% CL requires  $\sim 25,000$  samples. So is this 12,000 enough for the parameter estimation purposes? To test that we have also performed a full MCMC calculation from scratch (with 100,000 samples) using WMAP data along with a sound horizon measurement SH1. The results of this calculation are shown in Fig. 4.10. The contours in Fig. 4.8 although being noisier are very similar to the ones in Fig. 4.10. In fact the corresponding 1D distributions are practically indistinguishable. This shows that the initial use of importance sampling was indeed justified.



**Figure 4.9.:** The same as Fig. 4.8, only for the sound horizon measurement  $(103.0 \pm 7.6) h^{-1} \text{ Mpc}$  (SH2).



**Figure 4.10.:** The exact analog of Fig. 4.8, now for the full MCMC calculation.



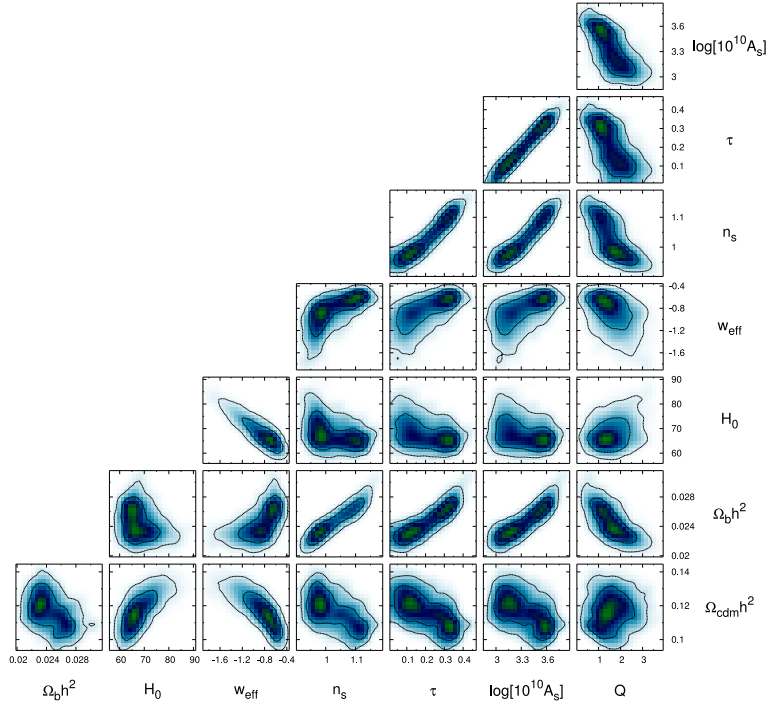
### 4.4.3. Constraints from the full power spectrum

Using the WMAP data and the SDSS LRG power spectrum as shown in Fig. 4.1 along with the power spectrum transformation and an additional new parameter  $Q$ , as described in Sect. 4.3, we build a 100,000 element Markov chain in the 8-dimensional parameter space. The resulting 2D parameter distribution functions are shown in Fig. 4.11. Here we see that in several cases distributions start to become doubly-peaked. Also the constraints on  $H_0$  and  $w_{\text{eff}}$  are weaker than the ones obtained in the previous subsection. On the other hand, now a rather strong constraint has been obtained for  $\Omega_{\text{cdm}}h^2$ . Even stronger constraint (not shown in the figure) is obtained for  $\Omega_m h$ — the shape parameter  $\Gamma$ . This just illustrates the well-known fact that the shape of the matter power spectrum is most sensitive to  $\Gamma$ . The new parameter  $Q$ , describing the deformation of the linear spectrum to the evolved redshift-space galaxy power spectrum, is seen to be significantly degenerate with several parameters e.g.  $\Omega_b h^2$ ,  $n_s$ ,  $\tau$ ,  $A_s$ . On the other hand it does not interfere too strongly with  $H_0$ .

It might seem strange that using the full data we obtain weaker constraints on  $H_0$  and  $w_{\text{eff}}$ . But after all, we should not be too surprised, since our understanding of how the linear spectrum is deformed to the evolved redshift-space power spectrum is rather limited. Here we were introducing an additional parameter  $Q$ , which starts to interfere with the rest of the parameter estimation. Also one should remind that maximum likelihood is the global fitting technique i.e. it is not very sensitive to specific features in the data. On the other hand, modeling of the oscillatory component of the spectrum does not call for any extra parameters. Also the underlying physics is much better understood. In fact, the observable low redshift acoustic scale is determined by four parameters only:  $\Omega_m h^2$ ,  $\Omega_b h^2$ ,  $w_{\text{eff}}$  and  $h$ . The optimal data analysis of course should incorporate both components: (i) general shape of the spectrum i.e. low frequency components, and (ii) oscillatory part, with appropriate weightings. It is clear that in the current “full spectrum” maximum likelihood analysis the acoustic features are weighted too weakly.

### 4.4.4. One dimensional distributions

To compare the measurements of the parameters in a more clear fashion we provide in Fig. 4.12 several 1D marginalized distributions. The 68% and 95% credible regions along with the medians of these distributions are provided in Table 4.1. Here the parameters  $\Omega_b h^2$ ,  $\Omega_{\text{cdm}} h^2$ ,  $\theta$ ,  $\tau$ ,  $w_{\text{eff}}$ ,  $n_s$ ,  $A_s$  and  $Q$  (the last in case of the full spectrum analysis only) are primary parameters as used in the MCMC calculations. All the rest:  $\Omega_\Lambda$ ,  $t_0$ ,  $\Omega_m$ ,  $z_{\text{reion}}$ ,  $H_0$ ,  $q_0$ ,  $j_0$  are derived from these. Here  $t_0$  is the age of the Universe,  $q_0$  the deceleration parameter and  $j_0$  the so-called jerk (see e.g. Blandford et al. 2005) at  $z = 0$ . From Fig. 4.12 we can see that many parameters stay essentially the same as determined by WMAP + HST data. On the other hand, a new precise measurement of  $H_0$ , thanks to the measurement of the low redshift sound horizon along with strong constraints on  $\Omega_b h^2$  and  $\Omega_{\text{cdm}} h^2$  from the CMB data, helps to determine  $\Omega_m$  (as well as  $\Omega_b$  and  $\Omega_{\text{cdm}}$  separately) rather precisely. The same applies to the case of the full spectrum analysis, which provides us with a good estimate for the shape parameter  $\Gamma = \Omega_m h$ . In both cases also the constraint on  $w_{\text{eff}}$  is significantly improved. New improved limits on  $\Omega_m$  and  $w_{\text{eff}}$  immediately transform to better constraints on  $q_0$  and  $j_0$  (see Eqs. (1.13), (1.14)). For the “vanilla”  $\Lambda$ CDM model with  $w = -1$  the jerk parameter  $j_0 = 1$ . We can see that at the moment jerk is still rather poorly constrained. Using only the observational data whose nature is very well understood, namely the CMB power



**Figure 4.11.:** The 2D marginalized distributions from the  $W_{\text{MAP}}$  + SDSS LRG full power spectrum MCMC calculation.

spectra along with the low redshift sound horizon measurement, we get very strong support for the accelerating Universe (i.e.  $q_0 < 0$ ). The values  $q_0 > 0$  are ruled out by  $1.4\sigma$ ,  $2.9\sigma$  and  $5.5\sigma$  in case of the  $W_{\text{MAP}}$  + HST,  $W_{\text{MAP}}$  + SDSS LRG SH2 and  $W_{\text{MAP}}$  + SDSS LRG SH1, respectively.

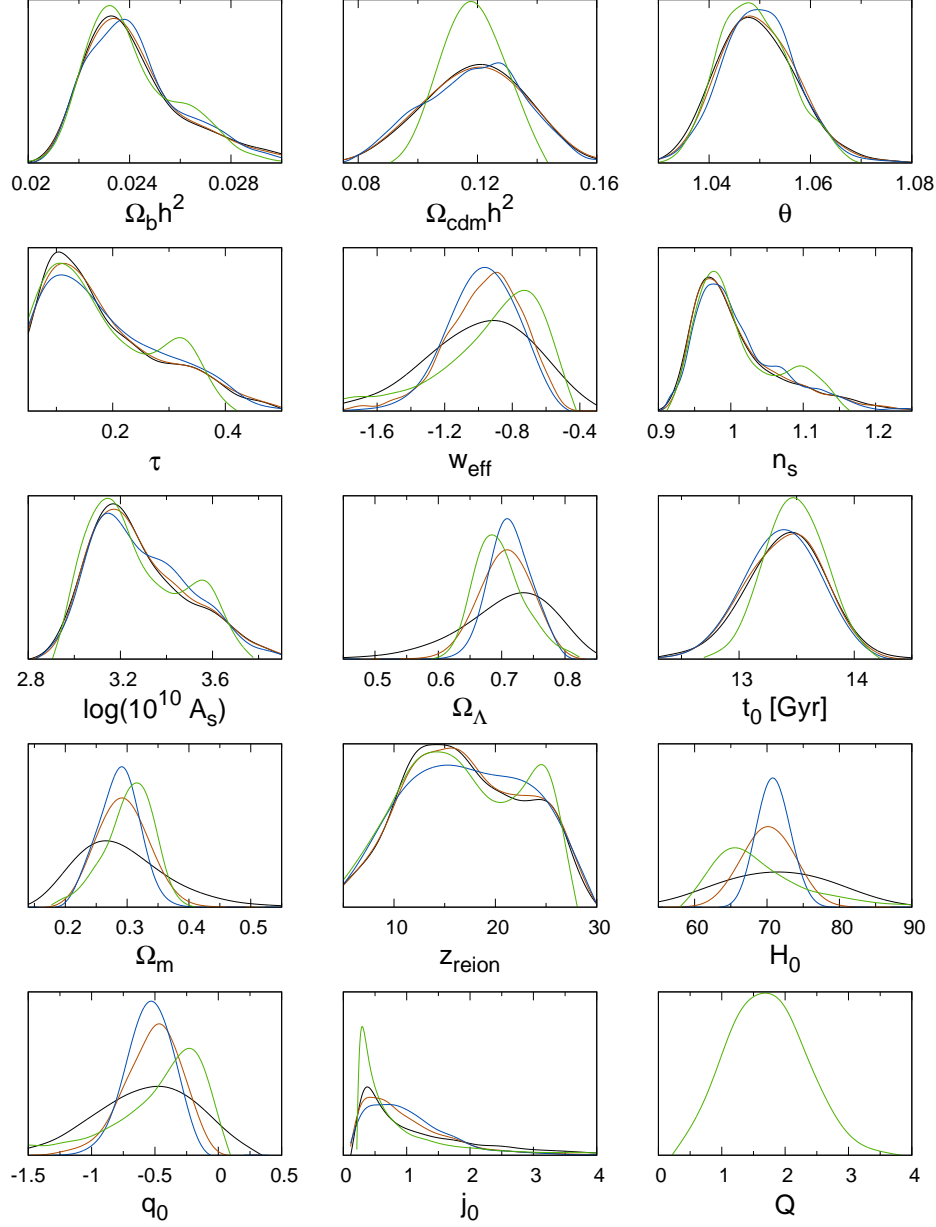
<sup>31</sup> Of course, one has to remind that current analysis assumed flat spatial sections.

#### 4.4.5. Most interesting constraints

We have shown that by adding the SDSS LRG clustering data to the  $W_{\text{MAP}}$  results we can get significantly tighter constraints on  $H_0$ ,  $\Omega_m$  and  $w_{\text{eff}}$  (or  $q_0$  and  $j_0$ ) than from the  $W_{\text{MAP}}$  + HST analysis alone. The comparison of the obtained limits on parameters  $H_0$ ,  $\Omega_m$  and  $w_{\text{eff}}$  is provided in Fig. 4.13. The largest error contours in both upper and lower group of panels correspond to  $W_{\text{MAP}}$  + HST, while the tightest to the  $W_{\text{MAP}}$  + SDSS LRG SH1 case. In the upper group of panels we have additionally given the constraints for the  $W_{\text{MAP}}$  + SDSS LRG SH2 case, whereas the lower group provides extra limits from the full spectrum +  $W_{\text{MAP}}$  analysis. In addition, in each of the panels we have given the degeneracy lines corresponding to the principal component given in Eq. (4.14). <sup>32</sup> In  $\Omega_m - H_0$  plane we have additionally plotted the lines corresponding to  $\Omega_m h^2 = \text{const}$  and  $\Gamma \equiv \Omega_m h = \text{const}$ . These are the combinations well determined by the CMB data and by the general shape of the matter power spectrum, respectively. As is evident from

<sup>31</sup>We can perform this analysis of the far away tails of the distributions since the  $W_{\text{MAP}}$  + HST chain we start with contains enough samples with  $q_0 > 0$  (see Fig. 4.12).

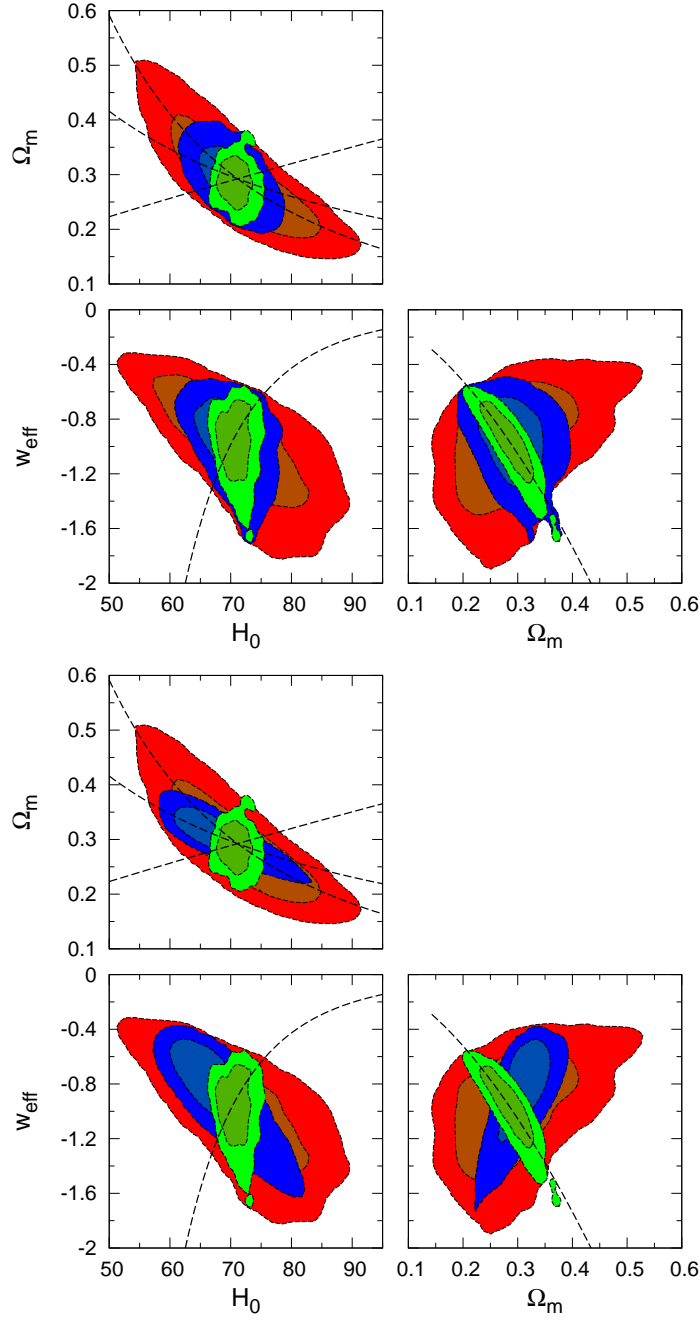
<sup>32</sup>The analog of Eq. (4.14) valid for the non-flat cases is given in Appendix A.



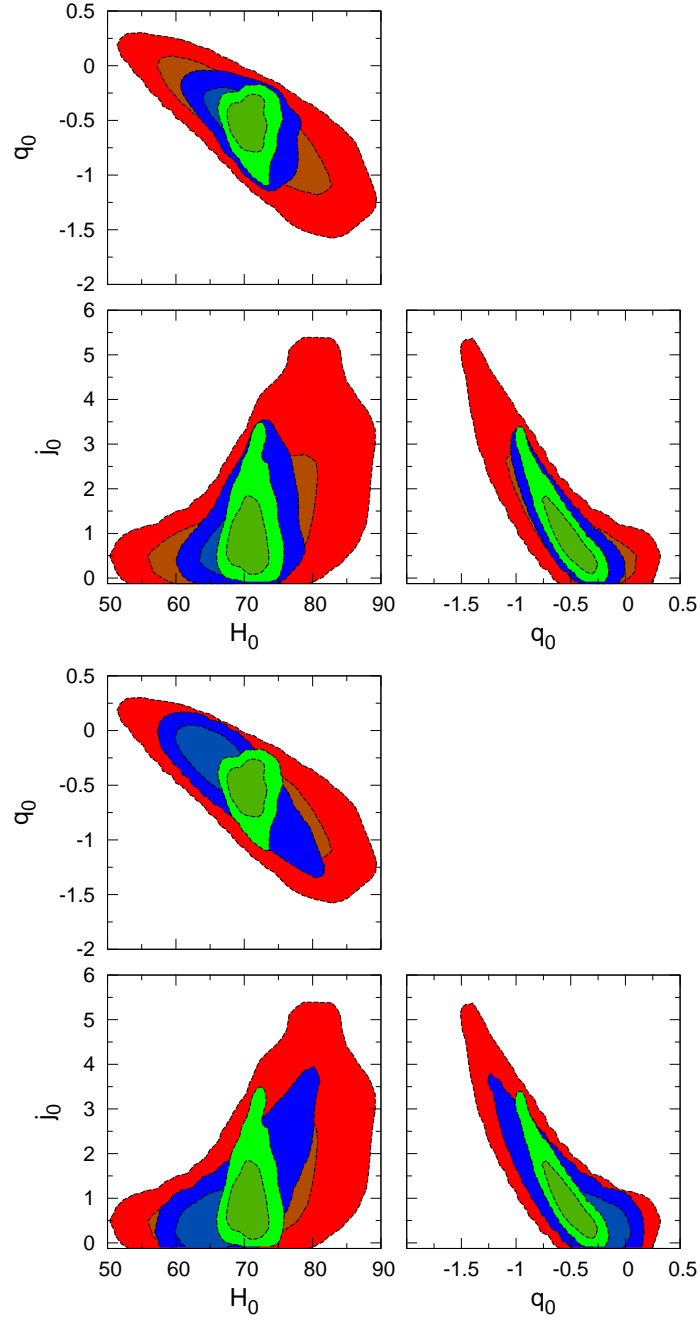
**Figure 4.12.:** The 1D posterior distributions for several cosmological parameters. Black, blue, red and green lines correspond to the  $W_{\text{MAP}} + \text{HST}$ ,  $W_{\text{MAP}} + \text{SDSS LRG SH1}$ ,  $W_{\text{MAP}} + \text{SDSS LRG SH2}$  and  $W_{\text{MAP}} + \text{SDSS LRG full power spectrum}$  cases, respectively. The compact summary of these results can be found in Table 4.1.

**Table 4.1.:** Various quantities of the 1D distributions shown in Fig. 4.12. The first group of parameters are the primary ones used in the MCMC calculations, the second group represents various derived quantities, and the last shows the parameters held fixed due to our prior assumptions. The last row of the table also gives the total number of free parameters, excluding the bias parameter that was marginalized out analytically, for all of the investigated cases.

Parameter	W <sub>MAP</sub> + HST			W <sub>MAP</sub> + SDSS LRG sound horizon, adiab. (SH1)			W <sub>MAP</sub> + SDSS LRG sound horizon (SH2)			W <sub>MAP</sub> + SDSS LRG full spectrum		
	Median	68%	95%	Median	68%	95%	Median	68%	95%	Median	68%	95%
$\Omega_b h^2$	0.0238	+0.0026 -0.0014	+0.0057 -0.0025	0.0239	+0.0024 -0.0016	+0.0051 -0.0025	0.0238	+0.0025 -0.0014	+0.0057 -0.0024	0.0237	+0.0025 -0.0013	+0.0043 -0.0023
$\Omega_{cdm} h^2$	0.120	+0.017 -0.018	+0.031 -0.033	0.120	+0.017 -0.019	+0.032 -0.033	0.120	+0.017 -0.018	+0.032 -0.034	0.118	+0.011 -0.010	+0.020 -0.018
$\theta$	1.0489	+0.0085 -0.0073	+0.0179 -0.0138	1.0498	+0.0069 -0.0069	+0.0155 -0.0136	1.0493	+0.0083 -0.0071	+0.0171 -0.0138	1.0486	+0.0075 -0.0065	+0.0146 -0.0119
$\tau$	0.153	+0.044 -0.053	+0.231 -0.153	0.160	+0.049 -0.160	+0.225 -0.160	0.156	+0.045 -0.156	+0.230 -0.156	0.152	+0.046 -0.152	+0.195 -0.152
$w_{\text{eff}}$	-0.97	+0.30 -0.36	+0.54 -0.73	-0.97	+0.21 -0.21	+0.32 -0.45	-0.94	+0.20 -0.25	+0.35 -0.55	-0.86	+0.21 -0.40	+0.34 -0.92
$n_s$	0.991	+0.084 -0.037	+0.183 -0.061	0.995	+0.077 -0.066	+0.162 -0.062	0.992	+0.082 -0.037	+0.184 -0.059	0.993	+0.092 -0.036	+0.141 -0.060
$\log(10^{10} A_s)$	3.24	+0.28 -0.16	+0.53 -0.28	3.25	+0.26 -0.17	+0.47 -0.30	3.24	+0.28 -0.17	+0.53 -0.28	3.22	+0.30 -0.16	+0.45 -0.24
$Q$	—	—	—	—	—	—	—	—	—	1.71	+0.63 -0.60	+1.25 -1.11
$H_0$	71.4	+8.0 -15.5	+14.9 -8.2	70.8	+1.9 -1.8	+3.7 -3.4	70.1	+3.6 -3.4	+6.6 -6.6	67.6	+7.7 -4.3	+17.9 -7.2
$q_0$	-0.54	+0.38 -0.45	+0.68 -0.89	-0.54	+0.17 -0.18	+0.29 -0.38	-0.50	+0.20 -0.24	+0.36 -0.51	-0.38	+0.24 -0.50	+0.38 -1.22
$j_0$	0.91	+1.55 -0.69	+4.11 -0.69	0.91	+0.75 -0.49	+1.90 -0.70	0.82	+0.87 -0.45	+2.36 -0.63	0.63	+1.43 -0.33	+5.31 -0.39
$\Omega_\Lambda$	0.717	+0.063 -0.083	+0.108 -0.189	0.712	+0.035 -0.028	+0.063 -0.055	0.708	+0.040 -0.041	+0.072 -0.082	0.693	+0.045 -0.034	+0.096 -0.064
$\Omega_m$	0.283	+0.083 -0.063	+0.189 -0.108	0.288	+0.028 -0.035	+0.055 -0.063	0.292	+0.040 -0.031	+0.072 -0.059	0.307	+0.034 -0.045	+0.064 -0.096
$t_0$	13.42	+0.31 -0.36	+0.64 -0.76	13.37	+0.31 -0.32	+0.59 -0.69	13.41	+0.31 -0.36	+0.59 -0.72	13.48	+0.27 -0.26	+0.53 -0.52
$z_{\text{reion}}$	16.5	+7.3 -5.5	+10.9 -10.3	17.0	+7.1 -6.3	+10.5 -10.3	16.8	+7.2 -5.6	+10.8 -10.5	16.4	+7.7 -5.9	+10.0 -10.6
$\Omega_k$	0	0	0	0	0	0	0	0	0	0	0	0
$\Omega_{\nu}$	0	0	0	0	0	0	0	0	0	0	0	0
$A_T$	0	0	0	0	0	0	0	0	0	0	0	0
$\frac{d \ln n_s}{d \ln k}$	0	0	0	0	0	0	0	0	0	0	0	0
	7 free parameters			7 free parameters			7 free parameters			8 free parameters		



**Figure 4.13.:** The comparison of constraints on  $H_0$ ,  $\Omega_m$  and  $w_{\text{eff}}$ . In all panels the largest error contours correspond to the WMAP + HST, while the tightest to the WMAP + SDSS LRG SH1 case. The upper group of panels shows additionally the constraints for the WMAP + SDSS LRG SH2 case, whereas the lower group provides extra limits from the full spectrum + WMAP analysis. The dashed lines in all the panels show the principal component from Eq. (4.14). The additional lines in  $\Omega_m - H_0$  plane provide the directions  $\Omega_m h^2 = \text{const}$  and  $\Gamma \equiv \Omega_m h = \text{const}$ .



**Figure 4.14.:** The version of Fig. 4.13 with the parameter trio  $(H_0, \Omega_m, w_{\text{eff}})$  replaced by  $(H_0, q_0, j_0)$ .

Fig. 4.13, the principal direction of the low redshift sound horizon constraint is always almost perpendicular to the corresponding  $W_{\text{MAP}} + \text{HST}$  error contours, demonstrating the high level of complementarity of this new measurement. For the spatially flat models with constant dark energy equation of state parameter there exists a unique relation between parameter pairs  $(\Omega_m, w_{\text{eff}})$  and  $(q_0, j_0)$  (see Eqs. (1.13), (1.14), (1.15), (1.16)). Fig. 4.14 presents similar plots to Fig. 4.13, now only for the parameter triad  $(H_0, q_0, j_0)$  instead. The parameters shown in Figs. 4.13 and 4.14 are the ones that determine the low redshift expansion law, which can be approximately expressed as in Eq. (1.11). Introducing the look-back time  $t_{\text{lb}} = t_0 - t$ , where  $t_0$  is the age of the Universe at  $z = 0$ , one can solve Eq. (1.11) for redshift:

$$z \simeq H_0 t_{\text{lb}} + \left(1 + \frac{q_0}{2}\right) H_0^2 t_{\text{lb}}^2 + \left(1 + q_0 + \frac{j_0}{6}\right) H_0^3 t_{\text{lb}}^3 + \dots \quad (4.15)$$

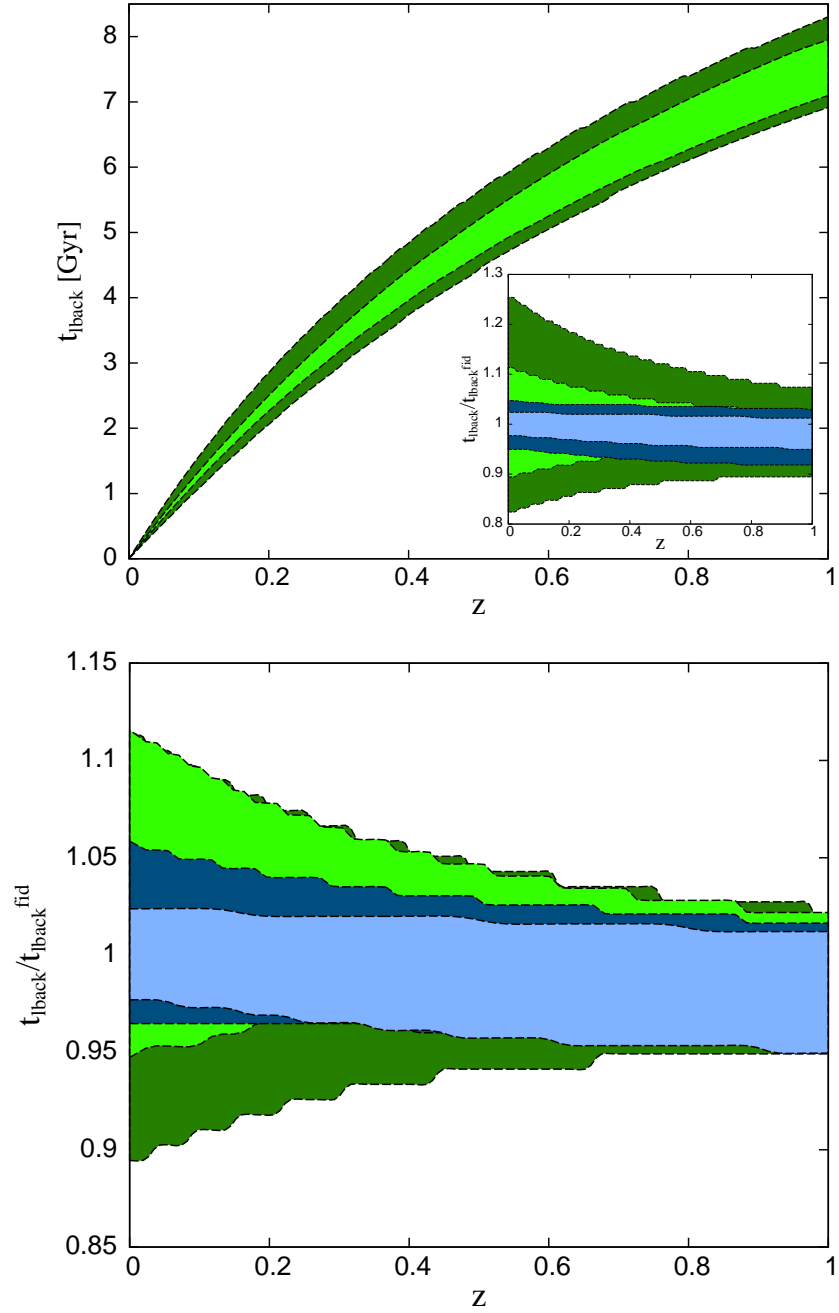
The precise calculation for the look-back time as a function of redshift is shown in Fig. 4.15. Here the upper panel shows the  $2\sigma$  regions corresponding to the  $W_{\text{MAP}} + \text{HST}$  and  $W_{\text{MAP}} + \text{SDSS LRG SH1}$ , respectively. The inset in the upper panel displays these regions after dividing by the look-back time corresponding to the best-fit  $W_{\text{MAP}}$  cosmology. Here in addition to the  $2\sigma$  contours also  $1\sigma$  regions are given. It is evident that the low redshift sound horizon measurement has helped to determine the recent expansion history of the Universe with much greater accuracy than available from the  $W_{\text{MAP}} + \text{HST}$  data alone. Of course, this is largely due to the much tighter constraint obtained for the Hubble parameter. The lower panel in Fig. 4.15 shows a similar plot than the inset in the upper panel. Here we have given only the  $1\sigma$  regions as a function of redshift for (starting from the bottommost) the  $W_{\text{MAP}} + \text{HST}$ ,  $W_{\text{MAP}} + \text{SDSS LRG full spectrum}$ ,  $W_{\text{MAP}} + \text{SDSS LRG SH2}$  and  $W_{\text{MAP}} + \text{SDSS LRG SH1}$  cases.

## 4.5. Discussion and Conclusions

In this paper we have performed a MCMC cosmological parameter study using the results from the recent SDSS LRG power spectrum analysis by [Hütsi \(2006a\)](#) along with the CMB temperature-temperature and temperature-polarization angular power spectra as determined by the  $W_{\text{MAP}}$  team ([Hinshaw et al. 2003](#), [Kogut et al. 2003](#)). We have carried out the analysis in two parts: (i) using the  $W_{\text{MAP}}$  data + the measurement of the low redshift sound horizon as found from the SDSS LRG redshift-space power spectrum, (ii) using the  $W_{\text{MAP}}$  data + full SDSS LRG power spectrum as shown in Fig. 4.1. As the formation of the acoustic features in the large-scale matter distribution is theoretically very well understood the separate treatment for the oscillatory part of the LRG power spectrum is well justified. Moreover, in comparison to the full power spectrum, which along with the dependence on several cosmological parameters requires additional modeling of the redshift-space distortions, nonlinear evolution, and biasing<sup>33</sup>, the acoustic scale depends on only a few cosmological parameters. The CMB measurements calibrate the physical scale of the sound horizon to a very good accuracy. By comparing it with the scale inferred from the low redshift LRG power spectrum measurements, we are able to get a very tight constraint on the Hubble parameter:  $h = 0.708^{+0.019}_{-0.018}$  if assuming adiabatic initial conditions, or  $h = 0.701^{+0.036}_{-0.034}$  if additional shift in oscillation phase is allowed. Having a tight

<sup>33</sup>Assuming we do not want to exclude the quasilinear scales from our analysis.





**Figure 4.15.:** Constraints on the low redshift expansion law. Upper panel:  $2\sigma$  credible regions of the look-back time as a function of redshift for the WMAP + HST and WMAP + SDSS LRG SH1 cases. The inset displays these regions after dividing by the look-back time corresponding to the best-fit WMAP cosmology. Here in addition to the  $2\sigma$  contours also  $1\sigma$  regions are given. Lower panel: Analog of the inset in the upper panel. Here we have given the  $1\sigma$  regions as a function of redshift for (starting from the bottommost) the WMAP + HST, WMAP + SDSS LRG full spectrum, WMAP + SDSS LRG SH2 and WMAP + SDSS LRG SH1 cases.

constraint on  $h$  allows us to break several parameter degeneracies, and thus helps us to determine various parameters like  $\Omega_m$ ,  $\Omega_{cdm}$ ,  $\Omega_b$  with a good precision. Also, in comparison to the WMAP + HST data, a significantly tighter constraint on  $w_{\text{eff}}$  is obtained. The full results for all the parameters are summarized in Table 4.1. The obtained values are in general in a good agreement with several other parameter studies e.g. Percival et al. (2002), Spergel et al. (2003), Tegmark et al. (2004b). Relatively tight bounds on  $(H_0, \Omega_m, w_{\text{eff}})$  or equivalently on  $(H_0, q_0, j_0)$  help us to determine the low redshift expansion law with significantly higher precision than available from the WMAP + HST data alone. If the initial fluctuations are constrained to be adiabatic, the measurement of the acoustic scale rules out a decelerating Universe, i.e.  $q_0 > 0$ , at  $5.5\sigma$  confidence level.

In contrast to the acoustic scale measurement, that gave a precise value for the Hubble parameter, the full spectrum provides us with a good estimate for the shape parameter  $\Gamma \equiv \Omega_m h = 0.207^{+0.011}_{-0.012}$ , which is in a very good agreement with the one found in Tegmark et al. (2004b). Since in the  $\Omega_m - h$  plane the  $\Gamma \equiv \Omega_m h = \text{const}$  line (see Fig. 4.13) is only relatively weakly tilted with respect to the relevant CMB degeneracy direction  $\Omega_m h^2 = \text{const}$ , the obtained limits on  $\Omega_m$  and  $h$  are not as strong as the ones obtained from the measurement of the acoustic scale. In contrast, the degeneracy lines corresponding to the low redshift acoustic scale measurement are in many cases almost orthogonal to the WMAP + HST “ellipses”, which explains the stronger constraints for several parameters.

In this paper we have also stressed the need to apply cosmological transformations to the theoretical model spectra before being compared with the relevant observational spectrum, which is valid only in the reference frame of the fiducial cosmological model that was used to analyze the data. So far almost all the parameter studies have completely ignored this point, which is probably fine for the shallow redshift surveys. On the other hand, in case of more deeper surveys like the SDSS LRG, reaching  $z \sim 0.5$ , these transformations have to be certainly applied. In general the line intervals along and perpendicular to the line of sight transform differently. Also the transformations depend on redshift. We have shown that for the samples like SDSS LRGs, with a typical redshift of  $z \sim 0.35$ , the single “isotropized” transformation taken at the median redshift of the survey provides a very accurate approximation to the more complete treatment.

For the parameter estimation we have used the SDSS LRG power spectrum down to the quasi-linear scales, which calls for the extra treatment of nonlinear effects, small scale redshift-space distortions and biasing. These additional complications can be relatively well dealt with the aid of the Halo Model (see Appendix D). We have shown in Hütsi (2006a) that a simple analytical model with additional four free parameters is able to approximate the observed spectrum to a very good precision. Also, the Halo Model has been shown to provide a good match to the results of the semianalytical galaxy formation studies (see e.g. Cooray & Sheth 2002). In this paper we have shown that to a rather tolerable accuracy the above four extra parameter Halo Model spectra (for reasonable values of the parameters) can be represented as a simple transformation of the linear power spectrum with only one extra parameter (see Fig. 4.2). The similar type of transformation was also used in Cole et al. (2005).

In order to investigate the possible biases introduced by the specific method used to extract the sound horizon from the power spectrum measurements, we have performed a Monte Carlo study, the results of what are shown in Fig. 4.7.

Finally, we remind that the current parameter study using the joint WMAP and SDSS LRG data assumed only spatially flat models. The parameter study including arbitrary curvature and also

possible massive neutrino component is planned for the future work.

## 5. Conclusions

To conclude we summarize the highlights of this thesis.

- We have studied the prospects of the upcoming wide-area blank-sky SZ cluster surveys, such as those to be performed by the PLANCK Surveyor spacecraft and the South Pole Telescope (SPT), to detect acoustic oscillations in the matter power spectrum. We have shown that at the largest scales these surveys are able to provide measurements of the matter power spectrum with better or similar accuracy <sup>1</sup> to the results obtainable from the SDSS Luminous Red Galaxy (LRG) survey, once completed within the next few years.
- Using the redshift-space cluster power spectra as obtainable by the PLANCK and the SPT SZ surveys, we have performed a Fisher matrix parameter forecasting study, particularly paying attention to the possibility for constraining the properties of dark energy. Although the obtainable constraints on the dark energy equation of state parameter are somewhat weaker than potentially obtainable through the study of the cluster number counts as a function of redshift, one has to keep in mind that the spatial clustering signal is significantly more robust to various systematic uncertainties. Our Fisher matrix calculations included the treatment for the anisotropic nature of the observed power spectrum due to redshift-space and cosmological distortions. The Halo Model used in these calculations was carefully calibrated against the outputs of the VIRGO Consortium's Hubble Volume simulations.
- The core part of this thesis was concerned with power spectrum analysis of the SDSS LRG sample. We have found evidence for a series of acoustic features in the LRG power spectrum down to scales of  $\sim 0.2 h \text{ Mpc}^{-1}$ , which corresponds to the 6th-7th peak in the CMB angular power spectrum. After correcting for nonlinearities and redshift-space distortions, the best-fit WMAP cosmological model was found to produce a very good match to the determined LRG power spectrum. This should be considered as another great success of the current cosmological “concordance” model.
- Under the assumption of adiabatic initial conditions and a distance-redshift relation given by the best-fit WMAP cosmology, the low redshift acoustic scale was measured to be  $(105.4 \pm 2.3) h^{-1} \text{ Mpc}$ . Using WMAP data together with the prior on the Hubble parameter from the HST Key Project,  $H_0 = 72 \pm 8 \text{ km/s/Mpc}$ , the corresponding scale would be predicted to be in the range  $(107 \pm 20) h^{-1} \text{ Mpc}$ , showing that our measurement provides approximately an order of magnitude improvement over that prediction.
- Using the obtained low redshift acoustic scale and also the full SDSS LRG power spectrum we have carried out the Maximum Likelihood cosmological parameter estimation via

---

<sup>1</sup>Assuming that cluster redshifts can be determined.

Markov Chain Monte Carlo techniques. In this analysis we focused on adiabatic, spatially flat models with negligible massive neutrino and tensor perturbation contributions. The simplest 6-parameter cosmological model was extended with the dark energy effective equation of state parameter  $w_{\text{eff}}$ . To break the parameter degeneracies additional data from the WMAP experiment was included. The most remarkable result is the constraint obtained for the Hubble parameter  $H_0 = 70.8^{+1.9}_{-1.8}$  km/s/Mpc. This precise measurement helped to break several parameter degeneracies and allowed us to measure the density parameters like  $\Omega_{\text{cdm}}$ ,  $\Omega_b$ , and also the dark energy equation of state parameter  $w_{\text{eff}}$  with significantly higher accuracy than available from the WMAP + HST data alone. Through the determination of these parameters we were able to constrain the low redshift expansion law of the Universe. Particularly, we found that a decelerating Universe is ruled out at the confidence level of  $5.5\sigma$ .

## A. Fitting formulae for the acoustic scales

The comoving distance traveled by the sound wave since the Big Bang up to redshift  $z$  can be expressed as:

$$s(z) = \int_{\infty}^z c_s(z')(1+z') \frac{dt}{dz'} dz', \quad (\text{A.1})$$

where the sound speed:

$$c_s(z) = \frac{c}{\sqrt{3[1+\mathcal{R}(z)]}}, \quad (\text{A.2})$$

$$\mathcal{R}(z) \equiv \frac{3\rho_b}{4\rho_\gamma} \simeq 3.04 \times 10^4 \cdot \frac{\Omega_b h^2}{z}. \quad (\text{A.3})$$

Using the Friedmann equation (see Eqs. (1.3),(1.4)) Eq. (A.1) can be integrated to yield (see e.g. Hu 1995)<sup>1</sup>:

$$s(z) = \frac{3.46 \times 10^3 \text{ Mpc}}{\sqrt{\Omega_m h^2 \cdot z_{\text{eq}} \mathcal{R}(z_{\text{eq}})}} \cdot \ln \left[ \frac{\sqrt{1+\mathcal{R}(z)} + \sqrt{\mathcal{R}(z) + \mathcal{R}(z_{\text{eq}})}}{1 + \sqrt{\mathcal{R}(z_{\text{eq}})}} \right], \quad (\text{A.4})$$

where the redshift for the matter-radiation equality can be given as:

$$z_{\text{eq}} \simeq 2.41 \times 10^4 \cdot \Omega_m h^2. \quad (\text{A.5})$$

The acoustic scale relevant for the CMB studies is  $s_* = s(z_*)$ , where  $z_*$  denotes the recombination redshift. For the “concordance” cosmological model the acoustic scale imprinted in the matter power spectrum  $s_d = s(z_d)$  is slightly larger than  $s_*$ . Here  $z_d$  denotes the redshift at which the baryons are released from the Compton drag of the photon field. To find accurate values for  $z_*$  and  $z_d$  one has to carry out a full calculation for the recombination history of the Universe. The results of these calculations can be conveniently expressed as the following fitting formulas (accurate at  $\sim 1\%$  level) (Hu & Sugiyama 1996, Eisenstein & Hu 1998):

$$z_* = 1048 \left[ 1 + 0.00124(\Omega_b h^2)^{-0.738} \right] \left[ 1 + g_1(\Omega_m h^2)^{g_2} \right], \quad (\text{A.6})$$

where

$$g_1 = 0.0783(\Omega_b h^2)^{-0.238} \left[ 1 + 39.5(\Omega_b h^2)^{0.763} \right]^{-1}, \quad (\text{A.7})$$

$$g_2 = 0.560 \left[ 1 + 21.1(\Omega_b h^2)^{1.81} \right]^{-1}, \quad (\text{A.8})$$

---

<sup>1</sup>The result is valid for high enough redshifts as relevant for the propagation of the sound waves.

and

$$z_d = 1291(\Omega_m h^2)^{0.251} \left[ 1 + 0.659(\Omega_m h^2)^{0.828} \right]^{-1} \left[ 1 + b_1(\Omega_b h^2)^{b_2} \right], \quad (\text{A.9})$$

where

$$b_1 = 0.313(\Omega_m h^2)^{-0.419} \left[ 1 + 0.607(\Omega_m h^2)^{0.674} \right], \quad (\text{A.10})$$

$$b_2 = 0.238(\Omega_m h^2)^{0.223}. \quad (\text{A.11})$$

For  $\Omega_b h^2 \lesssim 0.03$  the drag epoch follows the last scattering of the photons.

From the CMB measurements one can determine the angular scale  $\theta$  that corresponds to the sound horizon at decoupling with a good accuracy. This angle can be expressed as:

$$\theta = \frac{s_* h}{d_\perp(z_*)}, \quad (\text{A.12})$$

where  $s_* = s(z_*)$ , as given in Eq. (A.4), is measured in Mpc, and we have added an extra factor of  $h$  to the numerator to convert to the usual units of  $h^{-1}$  Mpc. Here  $d_\perp(z_*)$  (cf. Eq. (1.9)) is the comoving angular diameter distance to the last scattering surface, which is strongly dependent on the curvature radius  $R_0$  (cf. Eq. (1.6)). As  $s_*$  is only weakly dependent on  $\Omega_m h^2$  and  $\Omega_b h^2$ , it turns out that measurement of  $\theta$  is very sensitive to  $\Omega_k$ . The dependence of  $\theta$  on various cosmological parameters (around the “concordance” model point) is given in the upper panel of Fig. A.1. Using the following set of model parameters:  $(\Omega_m h^2, \Omega_b h^2, h, \Omega_{DE}, w_{\text{eff}})$ , the measurement of  $\theta$  constrains directly the linear combination  $0.40 \left( \frac{\Omega_m h^2}{0.14} \right) + 0.80 \left( \frac{\Omega_{DE}}{0.73} \right) - 0.45 \left( \frac{h}{0.71} \right)$ , or in case of logarithmic variables the combination  $(\Omega_m h^2)^{0.40} (\Omega_{DE})^{0.80} (h)^{-0.45}$ .

To measure the sound horizon  $s_d$  in the large-scale matter distribution one has to assume some background cosmological model in order to convert the observed redshifts to comoving distances. If the assumed fiducial model differs from the true cosmology, the measured scale will also be distorted. As shown in Chapter 4, for relatively low redshift measurements this distortion can be approximated as a single transformation for the “isotropized” comoving interval. (In general the comoving intervals along and perpendicular to the line of sight transform differently, and these transformations also depend on redshift.) The observed sound horizon  $s_d^{\text{obs}}$  can be approximated as:

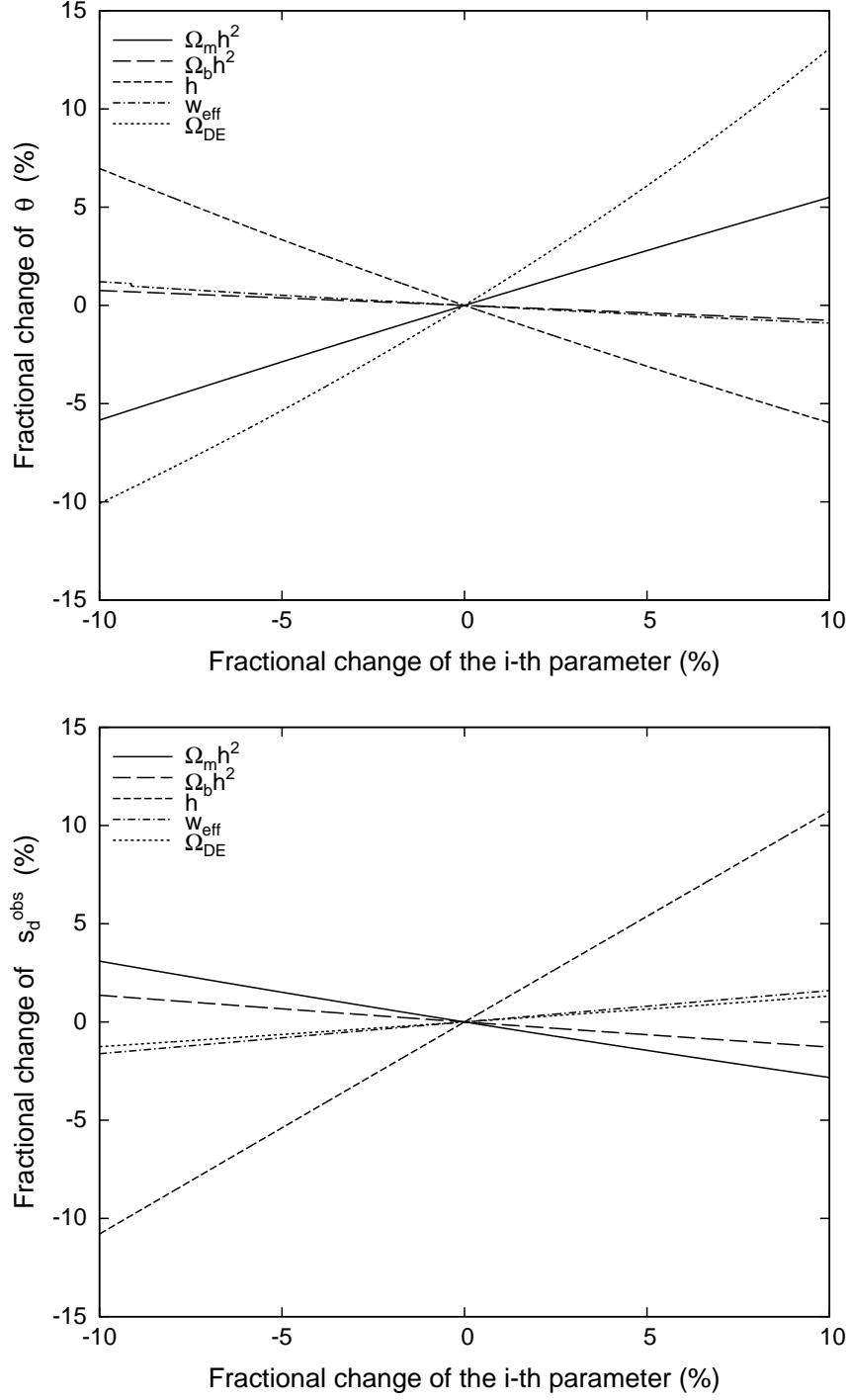
$$s_d^{\text{obs}} = c_{\text{isotr}} \cdot h s_d = \sqrt[3]{c_\parallel c_\perp^2} \cdot h s_d, \quad (\text{A.13})$$

where the extra factor of  $h$  is again included to convert to  $h^{-1}$  Mpc, and the functions  $c_\parallel(z)$ ,  $c_\perp(z)$ , which should be evaluated at the effective redshift  $z_{\text{eff}}$  of the observations (e.g. median redshift), are defined as:

$$c_\parallel(z) = \frac{E^{\text{fid}}(z)}{E(z)}, \quad c_\perp(z) = \frac{d_\perp(z)}{d_\perp^{\text{fid}}(z)}. \quad (\text{A.14})$$

The superscript  $^{\text{fid}}$  refers to the fiducial model and the functions  $E(z)$ ,  $d_\perp(z)$  are defined in Eqs. (1.4),(1.9). The dependence of  $s_d^{\text{obs}}$  on various cosmological parameters is shown in the lower panel of Fig. A.1. The fiducial model here was taken to be the best-fit WMAP cosmology and the “true models” were assumed to populate its intermediate neighborhood. We also assume  $z_{\text{eff}} = 0.35$  as in the case of the SDSS LRG sample analyzed in this thesis. Then the linear combination of parameters constrained by the measurement of  $s_d^{\text{obs}}$  turns out to be





**Figure A.1.:** The dependence of the angular scale corresponding to the sound horizon at decoupling (upper panel), and the sound horizon as measured from the low redshift matter power spectrum (lower panel), on various cosmological parameters. In the lower panel an effective observational redshift  $z_{\text{eff}} = 0.35$  has been assumed. The variation of the parameters has been performed around the “concordance” cosmological model. For this model the central values for  $\theta$  and  $s_d^{\text{obs}}$  are  $\sim 0.6^\circ$  and  $\sim 107 h^{-1} \text{ Mpc}$ , respectively.

$-0.26 \left( \frac{\Omega_m h^2}{0.14} \right) - 0.11 \left( \frac{\Omega_b h^2}{0.022} \right) + 0.94 \left( \frac{h}{0.71} \right) + 0.12 \left( \frac{\Omega_{DE}}{0.73} \right) + 0.15 \left( \frac{w_{\text{eff}}}{-1} \right)$ , or in case of logarithmic variables  $(\Omega_m h^2)^{-0.26} (\Omega_b h^2)^{-0.11} (h)^{0.94} (\Omega_{DE})^{0.12} (w_{\text{eff}})^{0.15}$ . Thus, as probably expected, the strongest dependence is on the Hubble parameter  $h$  i.e. the precise measurement of the acoustic scale at low redshifts should give us a good estimate for  $h$ .

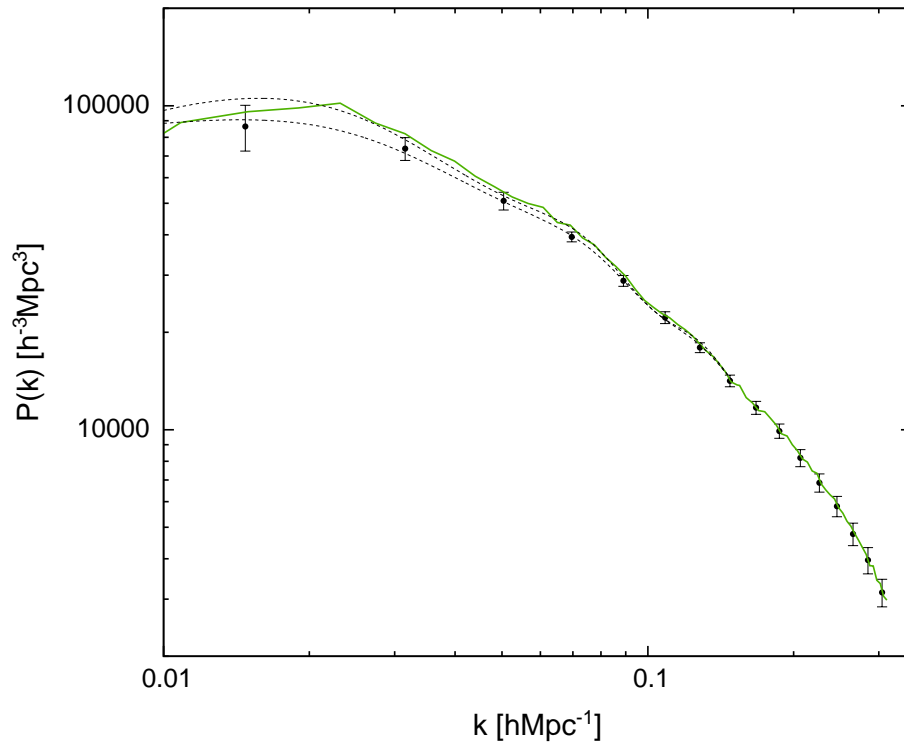
## B. Test problem

Here we present one test for our power spectrum calculation software<sup>1</sup>. As the input we use the  $z = 0$  cluster catalog of the VIRGO Hubble Volume simulations<sup>2</sup>, which covers the co-moving volume of  $3000 h^{-3} \text{ Mpc}^3$  and contains 1,560,995 clusters above the mass limit of  $6.75 \cdot 10^{13} h^{-1} M_{\odot}$ . The average bias parameter of this catalog is  $b = 1.9$ , which is comparable to the SDSS LRG value of  $b = 1.95$ . The power spectrum of the full sample is shown in Fig. B.1 with a solid line. Here for clarity we have not shown the errorbars, which are rather small for a sample of that size. Out of the full sample we generate 50 mock catalogs that have the same radial and angular selection functions as the SDSS LRG sample analyzed in this paper (see Figs. 3.2 and 3.3). The mean number of objects in the resulting catalogs is  $\sim 18,500$  i.e. the number density is roughly one third of the spatial density of the SDSS LRGs. Observer's location and pointing angles are taken randomly for each of the catalogs. The mean recovered power spectrum with  $1\sigma$  errorbars is shown in Fig. B.1. We see that the power spectrum of the underlying sample is recovered very well. On the largest scales there are some deviations, which can be explained as being caused by the smearing effect of the survey window. This is demonstrated by the dotted lines, where the lower/upper curve corresponds to the model spectrum with/without survey window convolution applied.

---

<sup>1</sup>Further tests can be found in [Hütsi \(2006b\)](#)

<sup>2</sup><http://www.mpa-garching.mpg.de/Virgo/>



**Figure B.1.:** Power spectra of galaxy clusters from the  $z = 0$  Hubble Volume simulation box. The solid line represents the spectrum for the full sample of 1,560,995 clusters. The circles with errorbars denote the recovered spectrum from the 50 mock catalogs having similar selection effects to the analyzed SDSS LRG sample. The dotted lines demonstrate the convolution effect of the survey window on the best fitting model spectrum.

## C. Mock catalogs

We build mock catalogs for the SDSS LRG by a 3 step procedure:

1. Generation of the density field using an optimized 2nd order Lagrangian perturbation calculation (2LPT).
2. Poisson sampling of the generated density field with the intensity of the process adjusted so, as to end up with a galaxy sample that has a clustering strength enhanced by a factor  $b^2$  with respect to the underlying field, and a number density equal to the observed LRG sample density at the minimal used redshift of 0.16 (see Fig. 3.2).
3. Extraction of the final catalog by applying the radial and angular selection function as given in Figs. 3.2 and 3.3, respectively.

In contrast to the Eulerian perturbation theory, where one does a perturbative expansion of the density contrast field, Lagrangian approach considers an expansion of the particle trajectories (see e.g. [Buchert & Ehlers 1993](#), [Bouchet et al. 1995](#), [Sahni & Coles 1995](#), [Bernardeau et al. 2002](#)). Here the central quantity is the displacement field  $\Psi(\mathbf{q})$ , which relates particle's initial comoving position (Lagrangian position)  $\mathbf{q}$  to its final Eulerian position  $\mathbf{x}$ :

$$\mathbf{x} = \mathbf{q} + \Psi(\mathbf{q}). \quad (\text{C.1})$$

It turns out that due to the decay of the rotational perturbation modes in the expanding Universe each order of the perturbation theory displacement field separates into a time-dependent and a Lagrangian position dependent factors ([Ehlers & Buchert 1997](#)). The position dependent part, due to its irrotational nature can be given as a gradient of a scalar potential. As a result, one can expand the displacement field as follows:

$$\Psi(\mathbf{q}) = D_1 \nabla_{\mathbf{q}} \phi^{(1)} + D_2 \nabla_{\mathbf{q}} \phi^{(2)}. \quad (\text{C.2})$$

Here the 1st term describes the classical Zeldovich approximation ([Zel'Dovich 1970](#)). The time independent potentials  $\phi^{(1)}$  and  $\phi^{(2)}$  are found from the Poisson equations:

$$\Delta \phi^{(1)}(\mathbf{q}) = -\delta(\mathbf{q}) \quad (\text{C.3})$$

and

$$\Delta \phi^{(2)}(\mathbf{q}) = \frac{1}{2} \sum_i \sum_j \left( \phi_{,ii}^{(1)}(\mathbf{q}) \phi_{,jj}^{(1)}(\mathbf{q}) - \phi_{,ij}^{(1)}(\mathbf{q}) \phi_{,ji}^{(1)}(\mathbf{q}) \right), \quad (\text{C.4})$$

where  $_{,i}$  denotes the partial derivative with respect to the Lagrangian coordinate  $q_i$ .  $\delta(\mathbf{q})$  is the initial density contrast. We generate  $\delta(\mathbf{q})$  using the standard Zeldovich approximation on a regular cubical grid.

$D_1$  in Eq. (C.2) is the linear growth factor. The second-order growth factor  $D_2$  for flat models with a cosmological constant is to a good precision approximated as (Bouchet et al. 1995):

$$D_2 \simeq -\frac{3}{7}\Omega_m^{-1/143}D_1^2. \quad (\text{C.5})$$

According to Eq. (C.1) and (C.2) the peculiar velocity field is given as:

$$\mathbf{v} = D_1 f_1 H \nabla_q \phi^{(1)} + D_2 f_2 H \nabla_q \phi^{(2)}. \quad (\text{C.6})$$

Here  $H \equiv \frac{\dot{a}}{a}$  and  $f_i \equiv \frac{d \ln D_i}{d \ln a}$ . For flat models with a cosmological constant logarithmic derivatives of the growth factors can be approximated as (Bouchet et al. 1995):

$$f_1 \simeq \Omega_m^{5/9}, \quad f_2 \simeq 2\Omega_m^{6/11}. \quad (\text{C.7})$$

Lagrangian perturbative approach works fine up to the 1st shell-crossing. After that the formed caustic structures will start to be wiped out, since the particles just keep on moving without noticing the gravitational pull of the dense sheets/filaments. It is possible to cure this problem significantly by filtering out the small-scale Fourier modes. This is what is meant by the “optimization”. The method applied to the 1st order Lagrangian perturbation calculation is known as the truncated Zeldovich approximation (e.g. Coles et al. 1993, Melott et al. 1994, Weiss et al. 1996). Weiss et al. (1996) suggest to remove the small-scale power by applying a Gaussian  $k$ -space filter with a characteristic smoothing scale  $k_{gs}$  to the initial density field. Thus the power spectrum of the filtered field is given by:

$$P_{\text{optimized}}(k) = P(k) \exp\left(-\frac{k^2}{k_{gs}^2}\right). \quad (\text{C.8})$$

They recommend the value  $k_{gs} \simeq 1.2k_{nl}$ , where the nonlinearity scale  $k_{nl}$  is defined as:

$$\frac{D_1^2}{(2\pi)^3} \int_0^{k_{nl}} d^3k P(k) = 1. \quad (\text{C.9})$$

Although they studied only models with  $\Omega_m = 1$ , it has been later shown by Hamana (1998) that this “recipe” performs well for arbitrary Friedmann-Lemaître-Robertson-Walker models.

In our calculations we assume the WMAP “concordance” cosmology (Spergel et al. 2003). Linear power spectrum is taken from Eisenstein & Hu (1998). We build 2LPT density field on a  $256^3$ -grid with  $5 h^{-1}$  Mpc cell size using the same number of particles as the number of grid cells.

<sup>1</sup> Four copies of this box are combined to form a bigger  $2560 \times 2560 \times 1280 h^{-3}$  Mpc<sup>3</sup> volume. Out of that big box a sample of “galaxies” is selected with a radial number density as given in Fig. 3.2 and with an angular mask presented in Fig. 3.3. The parameters of the Poisson sampler <sup>2</sup> are tuned to give a sample with a bias parameter  $b \simeq 2$  in agreement with the observed value for the SDSS LRG sample. The redshift-space catalog is built by altering the radial distances of

<sup>1</sup>Due to the rather big cell size the truncation of the initial spectrum has a rather mild effect.

<sup>2</sup>We use a simple model where the intensity of the inhomogeneous Poisson process is linearly related to the underlying density field.

---

the “galaxies” by  $\mathbf{v}_r/H_0$ , where  $\mathbf{v}_r$  is the radial component of the peculiar velocity field (see Eq. (C.6)) and  $H_0 = 100 h \text{ km/s/Mpc}$ .

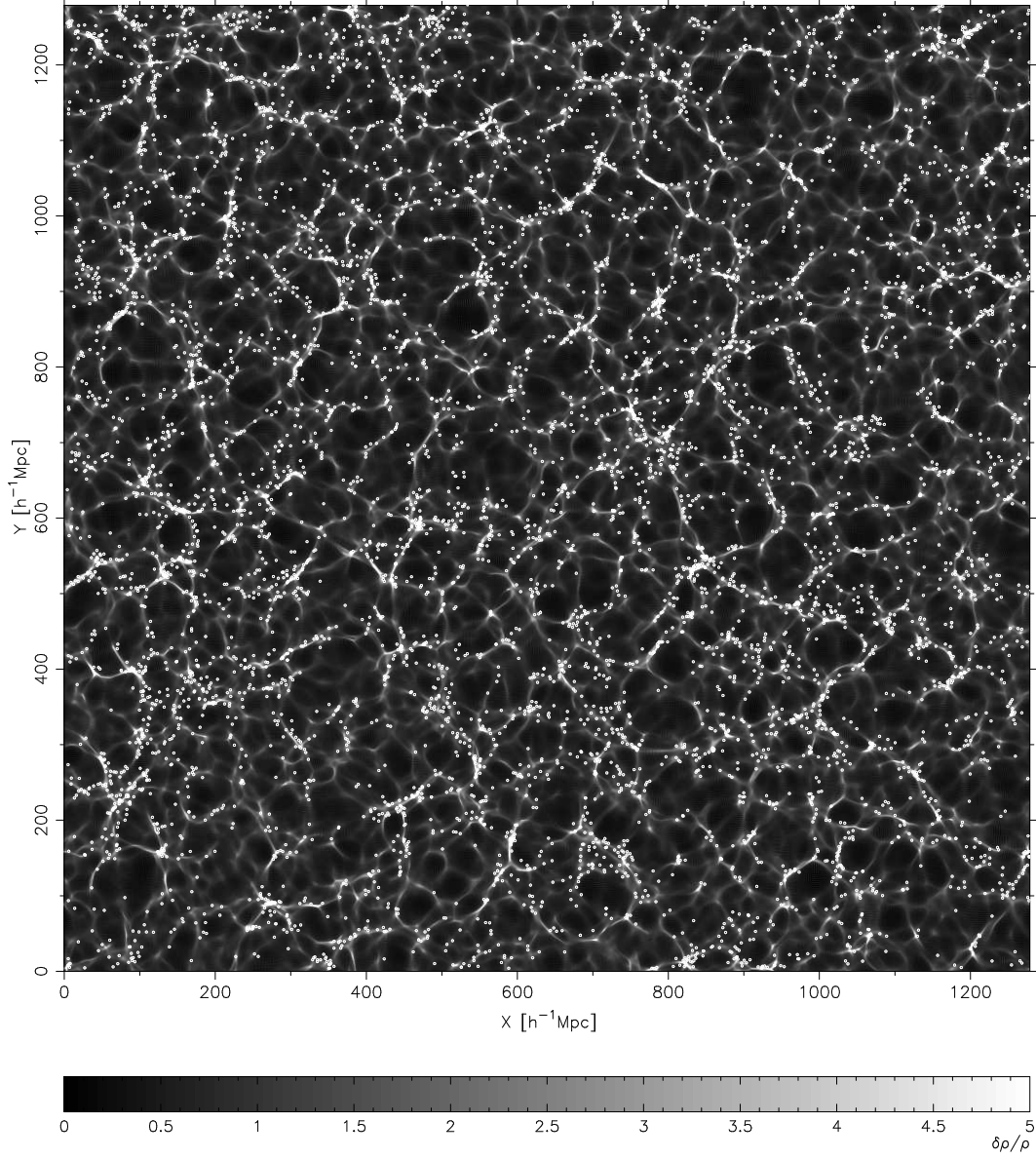
In Fig. C.1 we show a  $25 h^{-1} \text{ Mpc}$  thick slice through a box with  $1280 h^{-1} \text{ Mpc}$  side length. The underlying density field is presented as a gray scale image with white dots marking the positions of the “galaxies”. The power spectrum of the sample of  $\sim 350,000$  “galaxies” is shown in Fig. C.2<sup>3</sup>. We see that the shape of the spectrum is in good agreement with the linearly evolved power spectrum up to the scales of  $k \sim 0.5 h \text{ Mpc}^{-1}$ .

This approach gives us a “galaxy” sample that has rather realistic large-scale clustering properties. In terms of the Halo Model (see Appendix D) one can say that halo-halo clustering term is properly accounted for. 2LPT also gives reasonably accurate higher order correlations on quasi-nonlinear scales (e.g. Bouchet et al. 1995, Scoccimarro & Sheth 2002).

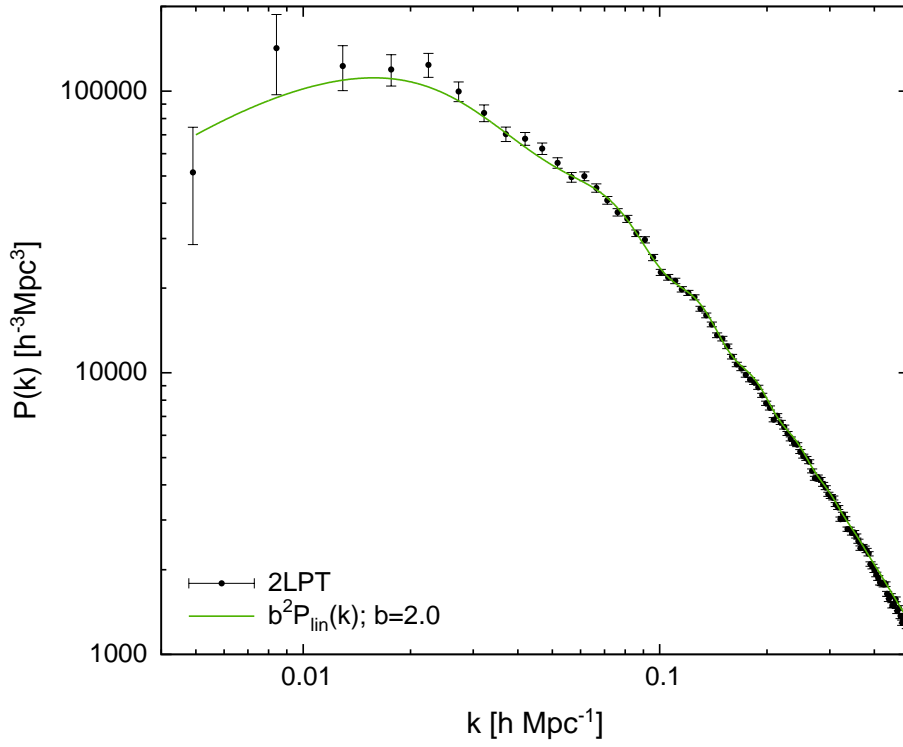
---

<sup>3</sup>Here as in the previous figure the SDSS LRG selection functions are not applied yet.





**Figure C.1.:** A  $25 h^{-1} \text{Mpc}$  thick slice through a  $1280 h^{-1} \text{Mpc}$  computational box. A gray scale image represents the underlying density field obtained by the optimized 2LPT approach. White dots mark the positions of the “galaxies” generated by the Poisson sampler.



**Figure C.2.:** The power spectrum of  $\sim 350,000$  “galaxies” from the simulation box shown in Fig. C.1. The solid line shows the linearly evolved input spectrum multiplied by the square of the bias parameter  $b = 2.0$ .



## D. Power spectrum from the halo model

The halo model description of the spatial clustering of galaxies is a development of the original idea by [Neyman & Scott \(1952\)](#), where one describes the correlations of the total point set as arising from the two separate terms: (i) 1-halo term, that describes the correlations of galaxies populating the same halo, (ii) 2-halo term, which takes into account correlations of the galaxies occupying different halos. For a thorough review see [Cooray & Sheth \(2002\)](#). Here we briefly give the results that are relevant to the current paper (see [Seljak 2001](#), [Cooray 2004](#)).

The power spectrum of galaxies in redshift space can be given as:

$$P(k) = P^{1h}(k) + P^{2h}(k), \quad (\text{D.1})$$

where the 1-halo term:

$$P^{1h}(k) = \int dM n(M) \frac{\langle N(N-1)|M \rangle}{\bar{n}^2} \mathcal{R}_p(k\sigma) |u_g(k|M)|^p, \quad (\text{D.2})$$

$$p = \begin{cases} 1 & \text{if } \langle N(N-1)|M \rangle < 1 \\ 2 & \text{if } \langle N(N-1)|M \rangle > 1 \end{cases} \quad (\text{D.3})$$

and the 2-halo term:

$$P^{2h}(k) = \left( \mathcal{F}_g^2 + \frac{2}{3} \mathcal{F}_v \mathcal{F}_g + \frac{1}{5} \mathcal{F}_v^2 \right) P_{\text{lin}}(k). \quad (\text{D.4})$$

Here:

$$\mathcal{R}_p \left( \alpha = k\sigma \sqrt{\frac{p}{2}} \right) = \frac{\sqrt{\pi} \text{erf}(\alpha)}{2} \frac{\alpha}{\alpha}, \quad (\text{D.5})$$

$$\mathcal{F}_g = \int dM n(M) b(M) \frac{\langle N|M \rangle}{\bar{n}} \mathcal{R}_1(k\sigma) u_g(k|M), \quad (\text{D.6})$$

$$\mathcal{F}_v = f \cdot \int dM n(M) b(M) \mathcal{R}_1(k\sigma) u(k|M). \quad (\text{D.7})$$

In the above expressions  $n(M)$  is the mass function and  $b(M)$  halo bias parameter. We calculate them using a prescription by [Sheth & Tormen \(1999\)](#) and [Sheth et al. \(2001\)](#).  $\bar{n}$  represents the mean number density of galaxies:

$$\bar{n} = \int dM n(M) \langle N|M \rangle. \quad (\text{D.8})$$

We take the mean of the halo occupation distribution in the following form:

$$\langle N|M \rangle = \left( \frac{M}{M_0} \right)^\alpha, \quad (\text{D.9})$$

where  $M_0$  and  $\alpha$  are free parameters. The second moment is chosen as (see Cooray 2004):

$$\langle N(N-1)|M \rangle = \beta^2(M) \langle N|M \rangle^2, \quad (\text{D.10})$$

$$\beta(M) = \begin{cases} \frac{1}{2} \log \left( \frac{M}{10^{11} h^{-1} M_\odot} \right) & \text{if } M < 10^{13} h^{-1} M_\odot \\ 1 & \text{otherwise.} \end{cases} \quad (\text{D.11})$$

$f$  in Eq. (D.7) denotes the logarithmic derivative of the linear growth factor:  $f \equiv \frac{d \ln D_1}{d \ln a}$ .  $u(k|M)$  and  $u_g(k|M)$  are the normalized Fourier transforms of the dark matter and galaxy density distributions within a halo of mass  $M$ . In our calculations we take both of these distributions given by the NFW profile (Navarro et al. 1997) and the concentration parameter  $c(M)$  is taken from Bullock et al. (2001). One dimensional velocity dispersion of the galaxies inside a halo with mass  $M$  is taken to follow the scaling of the isothermal sphere model:

$$\sigma = \gamma \sqrt{\frac{GM}{2R_{\text{vir}}}}, \quad (\text{D.12})$$

where  $R_{\text{vir}}$  is the virial radius of the halo and  $\gamma$  is a free parameter.

After specifying the background cosmology the above described model has four free parameters:  $M_0$ ,  $\alpha$  (Eq. (D.9)),  $\sigma$  (Eq. (D.12)) and  $M_{\text{low}}$ . The last parameter  $M_{\text{low}}$  represents the lower boundary of the mass integration i.e. halos with masses below  $M_{\text{low}}$  are assumed to be “dark”.

One can also use the halo model to estimate nonlinear contributions to the power spectrum covariance matrix. The additional term to the covariance matrix  $C_{ij}^{NL}$  ( $i, j$ —denote power spectrum bins) arising from the parallelogram configurations of the trispectrum<sup>1</sup> is given by (Cooray 2004):

$$C_{ij}^{NL} = \frac{T_{ij}}{V} = \frac{1}{V} \int_i \frac{d^3k}{V_i} \int_j \frac{d^3k}{V_j} \int dM n(M) \frac{\langle N(N-1)(N-2)(N-3)|M \rangle}{\bar{n}^4} |u_g(k_i|M)|^2 |u_g(k_j|M)|^2. \quad (\text{D.13})$$

$\int_i$  denotes an integral over a  $k$ -space shell centered at wavenumber  $k_i$  with a volume  $V_i = 4\pi k_i^2 \Delta k$ . The 4th moment of the halo occupation distribution is taken as:

$$\langle N(N-1)(N-2)(N-3)|M \rangle = \beta^2(M) \left[ 2\beta^2(M) - 1 \right] \left[ 3\beta^2(M) - 2 \right] \langle N|M \rangle^4, \quad (\text{D.14})$$

where  $\beta(M)$  and  $\langle N|M \rangle$  are given in Eqs. (D.11) and (D.9) above. Performing calculations in redshift space a factor of  $\mathcal{R}_p^2(k\sigma)$  (see Eq. (D.5)) must also be included in Eq. (D.13).

<sup>1</sup>Here only the contribution due to the 1-halo term is given.

## E. Fitting formulae for the coupling kernels

In this appendix we provide analytical fitting formulae for the coupling kernels  $K(k, k')$  in Eq. (3.24)<sup>1</sup>. The analytic form is motivated by the fact that the angle averaged survey window  $|W(k)|^2$  (see Fig. 3.8) can be reasonably well approximated by the analytical form:

$$|W(k)|^2 \equiv f(k) = \frac{1}{1 + \left(\frac{k}{a}\right)^2 + \left(\frac{k}{b}\right)^4}. \quad (\text{E.1})$$

Now assuming that  $|W(\mathbf{k})|^2$  is isotropic (which certainly is not the case as seen from Fig. 3.9), we can find the coupling kernels  $K(k, k')$  as:

$$K(k, k') = C \cdot \int d\Omega_{\mathbf{k}} \int d\Omega_{\mathbf{k}'} |W(\mathbf{k} - \mathbf{k}')|^2 = C \cdot \frac{8\pi^2}{kk'} \int_{|k-k'|}^{k+k'} f(x) x dx. \quad (\text{E.2})$$

For  $f(k)$  given by Eq. (E.1) the integral in Eq. (E.2) and the normalization constant  $C$  can be found analytically. The kernels are normalized such that

$$\int K(k, k') k'^2 dk' = 1 \quad (\text{E.3})$$

is satisfied.

Depending on the values of  $a$  and  $b$  there are two different solutions.

1.  $b^4 > 4a^4$ :

$$K(k, k') = \frac{C}{kk'} \ln \left[ \frac{g(k - k')}{g(k + k')} \right], \quad (\text{E.4})$$

where

$$g(x) = \frac{\mu + b^4 + 2a^2 x^2}{\mu - b^4 - 2a^2 x^2}, \quad (\text{E.5})$$

$$\mu = b^2 \sqrt{b^4 - 4a^4} \quad (\text{E.6})$$

and the normalization constant:

$$C = \frac{a}{\pi \sqrt{2} \left( \sqrt{b^4 + \mu} - \sqrt{b^4 - \mu} \right)}. \quad (\text{E.7})$$

---

<sup>1</sup>To avoid confusion we do not call them window functions since the word “window” has been already used to mean the Fourier transform of the survey volume.

2.  $b^4 < 4a^4$ :

$$K(k, k') = \frac{C}{kk'} [g(k + k') - g(k - k')] , \quad (\text{E.8})$$

where

$$g(x) = \arctan \left( \frac{b^4 + 2a^2 x^2}{b^2 \sqrt{4a^4 - b^4}} \right) \quad (\text{E.9})$$

and the normalization constant:

$$C = \frac{1}{\pi b \sqrt{2 - \left(\frac{b}{a}\right)^2}} . \quad (\text{E.10})$$

Although the isotropy assumption is certainly not correct, the above parametric family provides a very good fit to the numerically evaluated kernels as seen in Fig. 3.10. The best fitting  $a$  and  $b$  for the analyzed SDSS LRG sample are 0.00457 and 0.00475, respectively.

## F. Nonlinear model fitting. Correlated data

We find the best fitting parameters for the nonlinear model by minimizing  $\chi^2$ , which in the case of Gaussian errors is equivalent to finding the maximum likelihood solution. For this purpose we use Levenberg-Marquardt method as described in [Press et al. \(1992\)](#), where it was assumed that data values are uncorrelated. Since we are interested in the case with correlated errors, we have to make slight modifications to their implementation of the algorithm.

Using their notation,  $\chi^2$  is now calculated as:

$$\chi^2(\mathbf{a}) = \sum_{i=1}^N \sum_{j=1}^N [y_i - y(x_i; \mathbf{a})] \cdot C_{ij}^{-1} \cdot [y_j - y(x_j; \mathbf{a})], \quad (\text{F.1})$$

and the quantities  $\beta_k$  and  $\alpha_{kl}$  as follows:

$$\beta_k = \sum_{i=1}^N \sum_{j=1}^N [y_i - y(x_i; \mathbf{a})] \cdot C_{ij}^{-1} \cdot \frac{\partial y(x_j; \mathbf{a})}{\partial a_k}, \quad (\text{F.2})$$

$$\alpha_{kl} = \sum_{i=1}^N \sum_{j=1}^N \frac{\partial y(x_i; \mathbf{a})}{\partial a_k} \cdot C_{ij}^{-1} \cdot \frac{\partial y(x_j; \mathbf{a})}{\partial a_l}. \quad (\text{F.3})$$

In the above relations  $C_{ij}$  represents the data covariance matrix.





## G. Goodness of fit. Correlated Gaussian data

Under the assumption that statistical fluctuations  $\Delta y_i = y_i - y(x_i; \mathbf{a})$  ( $i = 1 \dots N$ ) in Eq. (F.1) are Gaussian distributed, with covariance matrix  $C_{ij}$ , one can easily derive probability density function (pdf) for the quantity  $\chi^2$ , and thus open up a way to estimate the goodness of fit.  $\chi^2$  goodness-of-fit estimator is usually exploited in the case of independent Gaussian variables. Here we show that calculating  $\chi^2$  for the correlated Gaussian data as given in Eq. (F.1), one obtains the same result that is well known for the independently distributed case.

According to our assumption  $\Delta \mathbf{y}$  is Gaussian distributed:

$$f_{\Delta \mathbf{y}}(\Delta \mathbf{y}) = \frac{1}{\sqrt{(2\pi)^N \det \mathbf{C}}} \exp\left(-\frac{1}{2} \Delta \mathbf{y}^T \cdot \mathbf{C}^{-1} \cdot \Delta \mathbf{y}\right). \quad (\text{G.1})$$

The conditional pdf of  $\chi^2$  given  $\Delta \mathbf{y}$ :

$$f_{\chi^2|\Delta \mathbf{y}}(\chi^2|\Delta \mathbf{y}) = \delta\left(\chi^2 - \Delta \mathbf{y}^T \cdot \mathbf{C}^{-1} \cdot \Delta \mathbf{y}\right), \quad (\text{G.2})$$

and so the pdf for  $\chi^2$  can be written as:

$$f_{\chi^2}(\chi^2) = \frac{1}{\sqrt{(2\pi)^N \det \mathbf{C}}} \int d^N \Delta \mathbf{y} \exp\left(-\frac{1}{2} \Delta \mathbf{y}^T \cdot \mathbf{C}^{-1} \cdot \Delta \mathbf{y}\right) \cdot \delta\left(\chi^2 - \Delta \mathbf{y}^T \cdot \mathbf{C}^{-1} \cdot \Delta \mathbf{y}\right). \quad (\text{G.3})$$

Now we define a new set of variables:

$$\Delta \mathbf{y}' = \mathbf{L}^T \cdot \Delta \mathbf{y}, \quad (\text{G.4})$$

where  $\mathbf{L}$  is the lower triangular matrix appearing in the Cholesky decomposition of  $\mathbf{C}^{-1}$ :

$$\mathbf{C}^{-1} = \mathbf{L} \cdot \mathbf{L}^T. \quad (\text{G.5})$$

Since  $\mathbf{C}^{-1}$  can be seen as the metric tensor, we can write for the transformation of the volume elements <sup>1</sup>:

$$\sqrt{\det \mathbf{C}^{-1}} d^N \Delta \mathbf{y} = \frac{d^N \Delta \mathbf{y}}{\sqrt{\det \mathbf{C}}} = d^N \Delta \mathbf{y}'. \quad (\text{G.6})$$

In the new frame, after changing to the spherical coordinates and integration over the angles:

$$f_{\chi^2}(\chi^2) = (2\pi)^{-\frac{N}{2}} \cdot \frac{\Omega_N}{2} \cdot \int d(\Delta y'^2) \Delta y'^{N-2} \exp\left(-\frac{\Delta y'^2}{2}\right) \cdot \delta\left(\chi^2 - \Delta y'^2\right), \quad (\text{G.7})$$

---

<sup>1</sup>The metric in the new frame is an identity matrix.

where the total  $N$ -dimensional solid angle:

$$\Omega_N = \frac{2\pi^{\frac{N}{2}}}{\Gamma\left(\frac{N}{2}\right)}. \quad (\text{G.8})$$

Thus the final result reads as:

$$f_{\chi^2}(\chi^2) = \frac{1}{2^{\frac{N}{2}}\Gamma\left(\frac{N}{2}\right)} (\chi^2)^{\frac{N}{2}-1} \exp\left(-\frac{\chi^2}{2}\right), \quad (\text{G.9})$$

which is a chi-square distribution with  $N$  degrees of freedom. Fitting  $P$  parameters (equivalent to adding  $P$  constraints), the effective number of degrees of freedom drops to  $N_{\text{eff}} = N - P$ , as usual.

Now it is straightforward to calculate  $p$ -values describing the goodness of fit.

## H. SDSS LRG power spectrum and covariance matrix

In Table H.1 we provide the measured SDSS LRG power spectrum  $\tilde{P}$  along with the covariance matrix from 1000 mock catalogs. We also give the corresponding deconvolved spectrum  $P$ . As the mode coupling kernels are relatively narrow the deconvolution problem can be solved rather satisfactorily. The deconvolved spectrum might be useful for the very fast parameter estimation applications. In case accurate Boltzmann solvers are used to calculate theoretical model spectra, the convolved spectrum is certainly preferred, since then the relative time taken for the extra convolution step is completely negligible.

We use the iterative deconvolution method due to Lucy (1974) with a specific implementation as given in Lin et al. (1996). The algorithm needs as an input the convolved power spectrum  $\tilde{P}_i$ , the wavenumbers  $k_i$  where the spectrum was measured, the description for the mode coupling kernels  $K(k, k')$  as defined in Eq. (3.25) (see Appendix E for the fitting formulae valid for the analyzed SDSS LRG sample.), and an initial guess for the deconvolved spectrum  $P_i^0$ . The index  $i = 1 \dots N$ , where  $N$  is the number of power spectrum bins. We take  $P_i^0 = \text{const} = 3 \times 10^4 h^{-3} \text{Mpc}^3$ . Having this information available we calculate:

$$K_{ij} = K(k_i, k_j) \quad (\text{H.1})$$

$$C_i = \sum_j K_{ij} \quad (\text{H.2})$$

$$D_i = k_i^2 \Delta k, \quad (\text{H.3})$$

where  $i, j = 1 \dots N$  and  $\Delta k$  is the wavenumber bin width, assumed here to be the same for all the bins.

Now the algorithm proceeds as follows:

1. Calculate  $A_i = D_i P_i^n$ .
2. Find  $\tilde{P}_i^n = \sum_j K_{ij} A_j$ .
3. Calculate  $B_i = \frac{\tilde{P}_i}{\tilde{P}_i^n}$ .
4. Find new estimate for the deconvolved power spectrum:  $P_i^{n+1} = \frac{P_i^n}{C_i} \sum_j K_{ij} B_j$ .
5. If not converged, go back to Step 1.

As an input we use the full 256-bin power spectrum. The resulting 256-bin deconvolved spectrum is thereafter rebinned to 16 bins.

**Table H.1.:** Measured SDSS LRG power spectrum  $\bar{P}$  and covariance matrix from 1000 mock catalogues. The corresponding deconvolved spectrum  $P$  is also given.

bin #	$k_l [h \text{ Mpc}^{-1}]$	$\bar{P} [h^{-3} \text{ Mpc}^3]$	$P [h^{-3} \text{ Mpc}^3]$	$\Delta P [h^{-3} \text{ Mpc}^3]$	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1	0.148e-1	0.133e6	0.172e6	0.228e5	0.518e9	0.358e8	0.762e7	0.521e7	0.337e7	0.161e7	0.189e7	0.987e6	0.128e7	0.106e7	0.400e6	0.236e6	0.204e6	0.114e6	0.106e6	0.184e6
2	0.315e-1	0.461e5	0.760e5	0.461e4	0.358e8	0.212e8	0.267e7	0.122e7	0.433e6	0.402e6	0.260e6	0.184e6	0.198e6	0.165e6	0.146e6	0.124e6	0.958e5	0.854e5	0.420e5	0.317e5
3	0.504e-1	0.500e5	0.500e5	0.224e4	0.762e7	0.267e7	0.500e7	0.103e7	0.362e6	0.316e6	0.192e6	0.170e6	0.810e5	0.702e5	0.515e5	0.467e5	0.389e5	0.285e5	0.152e5	0.138e5
4	0.697e-1	0.394e5	0.416e5	0.140e4	0.521e7	0.122e7	0.103e7	0.196e7	0.399e6	0.204e6	0.159e6	0.142e6	0.117e6	0.986e5	0.839e5	0.614e5	0.652e5	0.492e5	0.152e5	0.215e5
5	0.891e-1	0.286e5	0.294e5	0.816e3	0.337e7	0.432e6	0.362e6	0.399e6	0.667e6	0.160e6	0.112e6	0.110e6	0.788e5	0.654e5	0.537e5	0.477e5	0.356e5	0.256e5	0.188e5	0.146e5
6	0.109e0	0.213e5	0.214e5	0.550e3	0.161e7	0.402e6	0.316e6	0.204e6	0.160e6	0.303e6	0.113e6	0.798e5	0.614e5	0.520e5	0.416e5	0.427e5	0.325e5	0.235e5	0.166e5	0.123e5
7	0.128e0	0.185e5	0.191e5	0.462e3	0.189e7	0.260e6	0.192e6	0.159e6	0.112e6	0.113e6	0.213e6	0.817e5	0.594e5	0.436e5	0.384e5	0.284e5	0.229e5	0.256e5	0.188e5	0.146e5
8	0.148e0	0.147e5	0.150e5	0.364e3	0.987e6	0.184e6	0.170e6	0.142e6	0.110e6	0.798e5	0.817e5	0.133e6	0.533e5	0.366e5	0.312e5	0.269e5	0.217e5	0.188e5	0.181e5	0.129e5
9	0.167e0	0.119e5	0.118e5	0.295e3	0.128e7	0.198e6	0.810e5	0.117e6	0.788e5	0.614e5	0.594e5	0.533e5	0.871e5	0.384e5	0.266e5	0.242e5	0.217e5	0.188e5	0.188e5	0.129e5
10	0.187e0	0.103e5	0.105e5	0.251e3	0.106e7	0.165e6	0.702e5	0.986e5	0.654e5	0.520e5	0.456e5	0.360e5	0.384e5	0.631e5	0.254e5	0.200e5	0.169e5	0.136e5	0.136e5	0.100e5
11	0.206e0	0.869e4	0.866e4	0.224e3	0.400e6	0.146e6	0.515e5	0.839e5	0.537e5	0.416e5	0.384e5	0.312e5	0.266e5	0.254e5	0.504e5	0.237e5	0.181e5	0.158e5	0.114e5	0.123e5
12	0.226e0	0.759e4	0.764e4	0.210e3	0.246e6	0.124e6	0.467e5	0.614e5	0.647e5	0.427e5	0.284e5	0.269e5	0.242e5	0.200e5	0.237e5	0.441e5	0.217e5	0.174e5	0.148e5	0.141e5
13	0.246e0	0.654e4	0.648e4	0.199e3	0.204e6	0.958e5	0.389e5	0.652e5	0.564e5	0.422e5	0.299e5	0.217e5	0.188e5	0.169e5	0.181e5	0.217e5	0.396e5	0.191e5	0.133e5	0.146e5
14	0.265e0	0.599e4	0.605e4	0.182e3	0.114e6	0.854e5	0.285e5	0.492e5	0.477e5	0.325e5	0.256e5	0.188e5	0.188e5	0.136e5	0.158e5	0.174e5	0.191e5	0.330e5	0.152e5	0.120e5
15	0.285e0	0.530e4	0.531e4	0.162e3	0.106e6	0.420e5	0.138e5	0.152e5	0.266e5	0.233e5	0.215e5	0.181e5	0.128e5	0.927e4	0.114e5	0.148e5	0.133e5	0.152e5	0.261e5	0.155e5
16	0.305e0	0.479e4	0.479e4	0.164e3	0.184e6	0.317e5	0.151e5	0.215e5	0.298e5	0.265e5	0.217e5	0.129e5	0.143e5	0.100e5	0.123e5	0.141e5	0.146e5	0.120e5	0.155e5	0.271e5

# Bibliography

- Adelman-McCarthy, J. K., Agüeros, M. A., Allam, S. S., et al. 2005, ApJS, accepted; astro-ph/0507711
- Alcock, C. & Paczynski, B. 1979, Nature, 281, 358
- Bahcall, N. A., Ostriker, J. P., Perlmutter, S., & Steinhardt, P. J. 1999, Science, 284, 1481
- Bahcall, N. A. & Soneira, R. M. 1983, ApJ, 270, 20
- Bashinsky, S. & Bertschinger, E. 2002, Phys. Rev. D, 65, 123008
- Bennett, C. L., Halpern, M., Hinshaw, G., et al. 2003, ApJS, 148, 1
- Benoît, A., Ade, P., Amblard, A., et al. 2003, A&A, 399, L19
- Berlind, A. A., Weinberg, D. H., Benson, A. J., et al. 2003, ApJ, 593, 1
- Bernardeau, F., Colombi, S., Gaztañaga, E., & Scoccimarro, R. 2002, Physics Reports, 367, 1
- Bertschinger, E. 1995, astro-ph/9506070
- Blake, C. & Bridle, S. 2005, MNRAS, 366, 152
- Blake, C. & Glazebrook, K. 2003, ApJ, 594, 665
- Blanchard, A., Douspis, M., Rowan-Robinson, M., & Sarkar, S. 2003, A&A, 412, 35
- Blandford, R. D., Amin, M., Baltz, E. A., Mandel, K., & Marshall, P. J. 2005, in ASP Conf. Ser. 339: Observing Dark Energy
- Blanton, M. R., Lin, H., Lupton, R. H., et al. 2003, AJ, 125, 2276
- Blumenthal, G. R., Dekel, A., & Primack, J. R. 1988, ApJ, 326, 539
- Bouchet, F. R., Colombi, S., Hivon, E., & Juskiewicz, R. 1995, A&A, 296, 575
- Bouchet, F. R., Juskiewicz, R., Colombi, S., & Pellat, R. 1992, ApJL, 394, L5
- Bridle, S. L., Crittenden, R., Melchiorri, A., et al. 2002, MNRAS, 335, 1193
- Bryan, G. L. & Norman, M. L. 1998, ApJ, 495, 80
- Buchert, T. 1992, MNRAS, 254, 729

- Buchert, T. & Ehlers, J. 1993, MNRAS, 264, 375
- Bullock, J. S., Kolatt, T. S., Sigad, Y., et al. 2001, MNRAS, 321, 559
- Colberg, J. M., White, S. D. M., Yoshida, N., et al. 2000, MNRAS, 319, 209
- Cole, S., Percival, W. J., Peacock, J. A., et al. 2005, MNRAS, 362, 505
- Coles, P. 1993, MNRAS, 262, 1065
- Coles, P. & Lucchin, F. 1995, *Cosmology. The origin and evolution of cosmic structure* (Chichester: Wiley)
- Coles, P., Melott, A. L., & Shandarin, S. F. 1993, MNRAS, 260, 765
- Colless, M., Dalton, G., Maddox, S., et al. 2001, MNRAS, 328, 1039
- Cooray, A. 2004, MNRAS, 348, 250
- Cooray, A. & Sheth, R. 2002, *Physics Reports*, 372, 1
- de Lapparent, V., Geller, M. J., & Huchra, J. P. 1986, *ApJL*, 302, L1
- Dekel, A. & Lahav, O. 1999, *ApJ*, 520, 24
- Dodelson, S. 2003, *Modern cosmology* (Amsterdam: Academic Press)
- Donahue, M., Scharf, C. A., Mack, J., et al. 2002, *ApJ*, 569, 689
- Doran, M. 2005, *Journal of Cosmology and Astro-Particle Physics*, 10, 11
- Doroshkevich, A. G., Zel'Dovich, Y. B., & Syunyaev, R. A. 1978, *Soviet Astronomy*, 22, 523
- Durrer, R., Kunz, M., & Melchiorri, A. 2002, *Physics Reports*, 364, 1
- Ehlers, J. & Buchert, T. 1997, *General Relativity and Gravitation*, 29, 733
- Eisenstein, D. J., Annis, J., Gunn, J. E., et al. 2001, *AJ*, 122, 2267
- Eisenstein, D. J. & Hu, W. 1998, *ApJ*, 496, 605
- Eisenstein, D. J., Zehavi, I., Hogg, D. W., et al. 2005, *ApJ*, 633, 560
- Evrard, A. E., MacFarland, T. J., Couchman, H. M. P., et al. 2002, *ApJ*, 573, 7
- Feldman, H. A., Kaiser, N., & Peacock, J. A. 1994, *ApJ*, 426, 23
- Fixsen, D. J. & Mather, J. C. 2002, *ApJ*, 581, 817
- Freedman, W. L., Madore, B. F., Gibson, B. K., et al. 2001, *ApJ*, 553, 47
- Fry, J. N. 1984, *ApJ*, 279, 499

- Gilks, W. R., Richardson, S., & Spiegelhalter, D. J. 1996, Markov Chain Monte Carlo in practice (Boca Raton: Chapman & Hall/CRC)
- Grainge, K., Carreira, P., Cleary, K., et al. 2003, MNRAS, 341, L23
- Gurbatov, S. N., Saichev, A. I., & Shandarin, S. F. 1989, MNRAS, 236, 385
- Halverson, N. W., Leitch, E. M., Pryke, C., et al. 2002, ApJ, 568, 38
- Hamana, T. 1998, ApJL, 507, L1
- Hanany, S., Ade, P., Balbi, A., et al. 2000, ApJL, 545, L5
- Harrison, E. R. 1970, Phys. Rev. D, 1, 2726
- Hasinger, G. 2005, private communication
- Hastings, W. K. 1970, Biometrika, 57, 97
- Hawkins, E., Maddox, S., Cole, S., et al. 2003, MNRAS, 346, 78
- Hinshaw, G., Spergel, D. N., Verde, L., et al. 2003, ApJS, 148, 135
- Hockney, R. W. & Eastwood, J. W. 1988, Computer simulation using particles (Bristol: Hilger)
- Hu, W. 1995, PhD thesis, UC Berkeley
- Hu, W. 2002, Phys. Rev. D, 65, 023003
- Hu, W. & Haiman, Z. 2003, Phys. Rev. D, 68, 063004
- Hu, W. & Sugiyama, N. 1995, ApJ, 444, 489
- Hu, W. & Sugiyama, N. 1996, ApJ, 471, 542
- Huterer, D. & Starkman, G. 2003, Physical Review Letters, 90, 031301
- Hütsi, G. 2005, astro-ph/0507678
- Hütsi, G. 2006a, A&A, 449, 891
- Hütsi, G. 2006b, A&A, 446, 43
- Hütsi, G. 2006, A&A, submitted, astro-ph/0604129
- Jenkins, A., Frenk, C. S., White, S. D. M., et al. 2001, MNRAS, 321, 372
- Jimenez, R., Verde, L., Peiris, H., & Kosowsky, A. 2004, Phys. Rev. D, 70, 023005
- Jing, Y. P. 2005, ApJ, 620, 559
- Jungman, G., Kamionkowski, M., Kosowsky, A., & Spergel, D. N. 1996, Phys. Rev. D, 54, 1332



- Kaiser, N. 1984, *ApJL*, 284, L9
- Kaiser, N. 1987, *MNRAS*, 227, 1
- Kaplinghat, M., Knox, L., & Skordis, C. 2002, *ApJ*, 578, 665
- Knebe, A., Green, A., & Binney, J. 2001, *MNRAS*, 325, 845
- Kogut, A., Spergel, D. N., Barnes, C., et al. 2003, *ApJS*, 148, 161
- Kolb, E. W. & Turner, M. S. 1990, *The early universe* (MA: Addison-Wesley)
- Kosowsky, A., Milosavljevic, M., & Jimenez, R. 2002, *Phys. Rev. D*, 66, 063007
- Landy, S. D. & Szalay, A. S. 1993, *ApJ*, 412, 64
- Lewis, A. & Bridle, S. 2002, *Phys. Rev. D*, 66, 103511
- Lewis, A., Challinor, A., & Lasenby, A. 2000, *ApJ*, 538, 473
- Liddle, A. R. 2004, *MNRAS*, 351, L49
- Liddle, A. R. & Lyth, D. H. 2000, *Cosmological Inflation and Large-Scale Structure* (Cambridge: Cambridge University Press)
- Lin, H., Kirshner, R. P., Sheckman, S. A., et al. 1996, *ApJ*, 471, 617
- Linde, A. D. 1990, *Particle physics and inflationary cosmology* (Chur: Harwood Academic Publishers)
- Linder, E. V. 2003, *Phys. Rev. D*, 68, 083504
- Lucy, L. B. 1974, *AJ*, 79, 745
- Lupton, R. 1993, *Statistics in theory and practice* (Princeton, N.J.: Princeton University Press)
- Ma, C.-P. & Bertschinger, E. 1995, *ApJ*, 455, 7
- MacKay, D. J. C. 2003, *Information Theory, Inference and Learning Algorithms* (Cambridge: Cambridge University Press)
- Magira, H., Jing, Y. P., & Suto, Y. 2000, *ApJ*, 528, 30
- Majumdar, S. & Mohr, J. J. 2004, *ApJ*, 613, 41
- Martínez, V. J. & Saar, E. 2002, *Statistics of the Galaxy Distribution* (Boca Raton: Chapman & Hall/CRC)
- Matarrese, S., Coles, P., Lucchin, F., & Moscardini, L. 1997, *MNRAS*, 286, 115
- Meiksin, A., White, M., & Peacock, J. A. 1999, *MNRAS*, 304, 851
- Melott, A. L., Pellman, T. F., & Shandarin, S. F. 1994, *MNRAS*, 269, 626

- Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H., & Teller, E. 1953, *J. Chem. Phys.*, 21, 1087
- Mo, H. J. & White, S. D. M. 1996, *MNRAS*, 282, 347
- Moutarde, F., Alimi, J.-M., Bouchet, F. R., Pellat, R., & Ramani, A. 1991, *ApJ*, 382, 377
- Mukhanov, V. F., Feldman, H. A., & Brandenberger, R. H. 1992, *Physics Reports*, 215, 203
- Narayanan, V. K., Berlind, A. A., & Weinberg, D. H. 2000, *ApJ*, 528, 1
- Navarro, J. F., Frenk, C. S., & White, S. D. M. 1997, *ApJ*, 490, 493
- Neal, R. M. 1993, *Probabilistic Inference Using Markov Chain Monte Carlo Methods*, Tech. Rep. CRG-TR-91-1, Department of Computer Science. University of Toronto
- Neal, R. M. 2000, *Slice sampling*, Tech. Rep. No. 2005, Department of Computer Science. University of Toronto
- Netterfield, C. B., Ade, P. A. R., Bock, J. J., et al. 2002, *ApJ*, 571, 604
- Neyman, J. & Scott, E. L. 1952, *ApJ*, 116, 144
- Novosyadlyj, B., Durrer, R., & Lukash, V. N. 1999, *A&A*, 347, 799
- Padmanabhan, T. 1993, *Structure Formation in the Universe* (Cambridge: Cambridge University Press)
- Pearson, T. J., Mason, B. S., Readhead, A. C. S., et al. 2003, *ApJ*, 591, 556
- Peebles, P. J. E. 1980, *The large-scale structure of the universe* (Princeton, N.J.: Princeton University Press)
- Peebles, P. J. E. 1993, *Principles of physical cosmology* (Princeton, N.J.: Princeton University Press)
- Peebles, P. J. E. & Yu, J. T. 1970, *ApJ*, 162, 815
- Percival, W. J., Baugh, C. M., Bland-Hawthorn, J., et al. 2001, *MNRAS*, 327, 1297
- Percival, W. J., Sutherland, W., Peacock, J. A., et al. 2002, *MNRAS*, 337, 1068
- Percival, W. J., Verde, L., & Peacock, J. A. 2004, *MNRAS*, 347, 645
- Perlmutter, S., Aldering, G., Goldhaber, G., et al. 1999, *ApJ*, 517, 565
- Press, W. H. & Schechter, P. 1974, *ApJ*, 187, 425
- Press, W. H., Teukolsky, S. A., Vetterling, W. T., & Flannery, B. P. 1992, *Numerical recipes in FORTRAN. The art of scientific computing* (Cambridge: Cambridge University Press)

- Raftery, A. E. & Lewis, S. M. 1995, in *Practical Markov Chain Monte Carlo*. Edited by W.R. Gilks, D.J. Spiegelhalter, S. Richardson. London: Chapman & Hall
- Riess, A. G., Filippenko, A. V., Challis, P., et al. 1998, *AJ*, 116, 1009
- Ruhl, J., Ade, P. A. R., Carlstrom, J. E., et al. 2004, in *Astronomical Structures and Mechanisms Technology*. Edited by Antebi, Joseph; Lemke, Dietrich. *Proceedings of the SPIE*, Volume 5498, pp. 11-29 (2004)., 11–29
- Sahni, V. & Coles, P. 1995, *Physics Reports*, 262, 1
- Sandvik, H. B., Tegmark, M., Wang, X., & Zaldarriaga, M. 2004, *Phys. Rev. D*, 69, 063005
- Schaap, W. E. & van de Weygaert, R. 2000, *A&A*, 363, L29
- Scoccimarro, R. 2004, *Phys. Rev. D*, 70, 083007
- Scoccimarro, R. & Sheth, R. K. 2002, *MNRAS*, 329, 629
- Scott, P. F., Carreira, P., Cleary, K., et al. 2003, *MNRAS*, 341, 1076
- Seljak, U. 2000, *MNRAS*, 318, 203
- Seljak, U. 2001, *MNRAS*, 325, 1359
- Seljak, U. & Zaldarriaga, M. 1996, *ApJ*, 469, 437
- Seo, H. & Eisenstein, D. J. 2003, *ApJ*, 598, 720
- Seo, H.-J. & Eisenstein, D. J. 2005, *ApJ*, 633, 575
- Sheth, R. K., Mo, H. J., & Tormen, G. 2001, *MNRAS*, 323, 1
- Sheth, R. K. & Tormen, G. 1999, *MNRAS*, 308, 119
- Silk, J. 1968, *ApJ*, 151, 459
- Smoot, G. F., Bennett, C. L., Kogut, A., et al. 1992, *ApJL*, 396, L1
- Spergel, D. N., Verde, L., Peiris, H. V., et al. 2003, *ApJS*, 148, 175
- Springel, V., White, S. D. M., Jenkins, A., et al. 2005, *Nature*, 435, 629
- Stoughton, C., Lupton, R. H., Bernardi, M., et al. 2002, *AJ*, 123, 485
- Sunyaev, R. A. & Zeldovich, I. B. 1980, *ARA&A*, 18, 537
- Sunyaev, R. A. & Zeldovich, Y. B. 1970, *A&A Suppl.*, 7, 3
- Sunyaev, R. A. & Zeldovich, Y. B. 1972, *Comments on Astrophysics and Space Physics*, 4, 173
- Tegmark, M. 1997, *Physical Review Letters*, 79, 3806

- Tegmark, M., Blanton, M. R., Strauss, M. A., et al. 2004a, *ApJ*, 606, 702
- Tegmark, M., Hamilton, A. J. S., Strauss, M. A., Vogeley, M. S., & Szalay, A. S. 1998, *ApJ*, 499, 555
- Tegmark, M., Strauss, M. A., Blanton, M. R., et al. 2004b, *Phys. Rev. D*, 69, 103501
- Tegmark, M., Taylor, A. N., & Heavens, A. F. 1997, *ApJ*, 480, 22
- Verde, L., Peiris, H. V., Spergel, D. N., et al. 2003, *ApJS*, 148, 195
- Weinberg, S. 1972, *Gravitation and Cosmology: Principles and Applications of the General Theory of Relativity* (New York: John Wiley & Sons)
- Weiss, A. G., Gottlober, S., & Buchert, T. 1996, *MNRAS*, 278, 953
- White, M. 2005, *Astroparticle Physics*, 24, 334
- White, M., Hernquist, L., & Springel, V. 2002, *ApJ*, 579, 16
- Yamamoto, K., Nishioka, H., & Suto, Y. 1999, *ApJ*, 527, 488
- Yamamoto, K. & Suto, Y. 1999, *ApJ*, 517, 1
- York, D. G., Adelman, J., Anderson, J. E., et al. 2000, *AJ*, 120, 1579
- Zaroubi, S. & Hoffman, Y. 1996, *ApJ*, 462, 25
- Zehavi, I., Eisenstein, D. J., Nichol, R. C., et al. 2005, *ApJ*, 621, 22
- Zel'Dovich, Y. B. 1970, *A&A*, 5, 84
- Zeldovich, Y. B. 1972, *MNRAS*, 160, 1P



# Acknowledgments

First of all I would like to thank my supervisor Prof. Rashid Sunyaev for making it possible for me to complete my PhD studies in such an outstanding institute as MPA, and also for suggesting a very promising topic to work with. In the beginning I did not believe myself that it might be possible to detect baryonic features in the large-scale matter distribution, but after a while I had to change my initial opinion.

My research in MPA has greatly benefitted from fruitful discussions with several colleagues, most importantly with Jens Chluba, Carlos Hernández-Monteagudo, Jose Alberto Rubiño-Martín and Dimitrios Giannios. Thank you!

Even PhD students find free time now and then. Here I would like to use an opportunity to thank the members of our great hiking team: Jens Chluba, Carlos aka Pablo Hernández-Monteagudo and Rasmus the Viking Voss. Many thanks also to the members of the Munich beer fan club, including my office mate and the “biggest mouth” at MPA Herrn Dr. Maximilian Stritzinger and the “Caveman” Jonathan Braithwaite, who are also acknowledged for their great language editing skills.

Also I am very thankful to the collaborators from the Cosmology Group of my home institute in Estonia, especially Jaan Einasto, Enn Saar and Mirt Gramann.

And finally, I would like to thank my parents and my brother for the constant support and encouragement throughout all the years.

Some of the N-body simulations used in this thesis were carried out by the Virgo Supercomputing Consortium using computers based at the Computing Centre of the Max-Planck Society in Garching and at the Edinburgh parallel Computing Centre. The data are publicly available at <http://www.mpa-garching.mpg.de/NumCos>.

Funding for the creation and distribution of the SDSS Archive has been provided by the Alfred P. Sloan Foundation, the Participating Institutions, the National Aeronautics and Space Administration, the National Science Foundation, the U.S. Department of Energy, the Japanese Monbukagakusho, and the Max Planck Society. The SDSS Web site is <http://www.sdss.org/>.

The SDSS is managed by the Astrophysical Research Consortium (ARC) for the Participating Institutions. The Participating Institutions are The University of Chicago, Fermilab, the Institute for Advanced Study, the Japan Participation Group, The Johns Hopkins University, the Korean Scientist Group, Los Alamos National Laboratory, the Max-Planck-Institute for Astronomy (MPIA), the Max-Planck-Institute for Astrophysics (MPA), New Mexico State University, University of Pittsburgh, University of Portsmouth, Princeton University, the United States Naval Observatory, and the University of Washington.



# CURRICULUM VITAE AND LIST OF PUBLICATIONS

GERT HÜTSI

Max-Planck-Institut für Astrophysik      Tel: +49 89 30000 2190  
Karl-Schwarzschild-Str. 1      Fax: +49 89 30000 2235  
D-85741 Garching, Germany      gert@mpa-garching.mpg.de

## PERSONAL DATA

---

Date of birth: 3 January 1975  
Place of birth: Tallinn, Estonia  
Nationality: Estonian

## EDUCATION

---

- |                        |   |
|------------------------|---|
| Sept. 2002 – present   | Research for Ph.D. at the <i>Max-Planck-Institut für Astrophysik</i><br>Supervisor: Prof. Rashid Sunyaev<br>Thesis to be defended in spring 2006  |
| Sept. 1998 – June 2000 | M.Sc. in Theoretical Physics at the <i>University of Tartu</i> (Estonia)<br>Thesis title: “Models with a steplike initial spectrum and large-scale structure of the Universe”<br>Supervisor: Dr. Mirt Gramann         |
| Jan. 1996 – July 1996  | study and research at the <i>Uppsala University</i> (Sweden) as an exchange student   |
| Sept. 1994 – June 1998 | Diploma (B.Sc.) in Physics at the <i>University of Tartu</i> (Estonia)<br>Thesis title: “The mass function and peculiar velocities of galaxy clusters in various cosmological models”<br>Supervisor: Dr. Mirt Gramann |

## PUBLICATIONS

---

### *Refereed Publications*

- Acoustic oscillations in the SDSS DR4 Luminous Red Galaxy sample power spectrum  
**Hütsi, G.**  
*Astronomy & Astrophysics* 449, 891 (2006)



- Clustering of SZ clusters on a past light-cone: acoustic oscillations and constraints on dark energy  
**Hütsi, G.**  
*Astronomy & Astrophysics* 446, 43 (2006)
- Clusters of galaxies in the microwave band: influence of the motion of the Solar System  
Chluba, J., **Hütsi, G.**, Sunyaev, R.A.  
*Astronomy & Astrophysics* 434, 811 (2005)
- Clusters and superclusters in the Las Campanas redshift survey  
Einasto, J., Einasto, M., **Hütsi, G.** et al.  
*Astronomy & Astrophysics* 410, 425 (2003)
- Clusters and superclusters in the Sloan Digital Sky Survey  
Einasto, J., **Hütsi, G.**, Einasto, M. et al.  
*Astronomy & Astrophysics* 405, 425 (2003)
- Primordial feature at the scale of superclusters of galaxies  
Gramann, M., **Hütsi, G.**  
*MNRAS* 327, 538 (2001)
- CDM models with a steplike initial power spectrum  
Gramann, M., **Hütsi, G.**  
*MNRAS* 316, 631 (2000)

*In preparation*

- Power spectrum of the SDSS Luminous Red Galaxies: constraints on cosmological parameters  
**Hütsi, G.**  
*Astronomy & Astrophysics*, submitted, astro-ph/0604129